Estimação e modelagem de taxas específicas de migração por idade e sexo

Ernesto F. L. Amaral

Texas A&M University

amaral@tamu.edu

Bernardo Lanza Queiroz

Universidade Federal de Minas Gerais (UFMG)

lanza@cedeplar.ufmg.br

Resumo

A importância da estimação de taxas específicas de migração por idade e sexo pode ser comparada àquela das Taxas Específicas de Fecundidade e de Mortalidade. O objetivo do capítulo é de ilustrar como essas taxas são estimadas para migração através de diferentes modelos, assim como discutir sua utilidade para analisar o nível e padrão da migração. Como há grande variação das taxas por idade, também é ressaltada a importância da modelagem das taxas de migração proposta por Rogers e Castro (1981) e seus avanços com o modelo log-linear de Rogers et al. (2002). Em relação aos migrantes recentes, que podem ser sub-enumerados pelo censo, uma alternativa metodológica é a de Xu-Doeve (2005; 2007). A modelagem das taxas de migração permite calcular uma série de parâmetros que podem ser interpretados separadamente para entender o nível e padrão da migração. As diferentes formulações matemáticas são ainda úteis para o estudo de migração em populações com dados de baixa qualidade, assim como para movimentos migratórios com baixos níveis (migração de retorno, idosos...).

Versão Preliminar

Documento preparado para o livro de técnicas de migração organizado pelo Professor José Irineu Rigotti da Universidade Federal de Minas Gerais (UFMG).

Citação

Amaral EFL, Queiroz BL. 2014. “Estimação e modelagem de taxas específicas de migração por idade e sexo.” *Open Science Framework Preprints*, September 4.

(<http://doi.org/10.31219/osf.io/z7gak>)

1. Introdução

A importância da estimação de taxas específicas de migração por idade e sexo pode ser comparada àquela das Taxas Específicas de Fecundidade e de Mortalidade. O objetivo do capítulo é de ilustrar como essas taxas são estimadas para migração através de diferentes modelos, assim como discutir sua utilidade para analisar o nível e padrão da migração. Os métodos desenvolvidos podem ser aplicados a diferentes países e bancos de dados (censos ou pesquisas domiciliares) para gerar estimativas de taxas de migração utilizando informações sobre o local de residência anterior (dados de última etapa), informação sobre migração de data-fixa e duração de residência no local de destino. Os resultados sugerem que os modelos funcionam muito bem em diferentes cenários o que abre enormes possibilidades de estudos de migração para pequenas áreas ou países com dados restritos.

Como há grande variação das taxas por idade, também é ressaltada a importância da modelagem das taxas de migração proposta por Rogers e Castro (1981) e seus avanços com o modelo log-linear de Rogers et al. (2002). Em relação aos migrantes recentes, que podem ser sub-enumerados pelo censo, uma alternativa metodológica é a de Xu-Doeve (2005; 2007). A modelagem das taxas de migração permite calcular uma série de parâmetros que podem ser interpretados separadamente para entender o nível e padrão da migração. As diferentes formulações matemáticas são ainda úteis para o estudo de migração em populações com dados de baixa qualidade, assim como para movimentos migratórios com baixos níveis (migração de retorno, idosos, entre outros sub-grupos). Utilizamos exemplos do Brasil, baseados nos censos de 1991 e 2000; do México, com base nos Censos de 1990 e 2000 e da África do Sul, Censo de 1996, para ilustrar os métodos apresentados no capítulo.

2. Modelando o padrão etário da migração – Rogers e Castro

Com base em estimativas de taxas de migração por idade, é possível aplicar os modelos matemáticos por Rogers e Castro (1981) para obter o ajuste da curva de migração. A metodologia para modelagem de funções de migração em um conjunto de 17 países. A justificativa para a elaboração desses modelos é que os comportamentos por idade da fecundidade, da mortalidade e da migração variam dentro de limites predeterminados, para a maioria das populações humanas.

A modelagem das taxas é uma tarefa complexa, em virtude da definição dos parâmetros iniciais e variabilidade amostral. Essa variação ocorre porque os dados de migração são captados por

amostragem, e porque a migração é um evento raro que não ocorre com a mesma frequência que a fecundidade e a mortalidade. Essas dificuldades ocorrem também por questões substantivas, relacionadas ao deslocamento da migração da esfera do mercado de trabalho e às limitações do modelo para se ajustar aos novos condicionantes.

O modelo proposto se baseia nas regularidades encontradas nos perfis etários da migração, o que permite estimar curvas de migração hipotéticas que podem ser usadas em estudos demográficos quando os dados são limitados, inadequados ou se está trabalhando com pequenas áreas. Em outros casos, os modelos matemáticos são úteis para suavizar taxas de migração calculadas diretamente dos dados, mas que podem ser bastante influenciadas por outliers ou quando não se tem informações suficientes sobre o padrão migratório de uma população. Com base nos modelos matemáticos é possível obter informações adequadas sobre as taxas de migração e isso auxilia no entendimento da mobilidade espacial das populações entre diferentes áreas.

O modelo proposto, assim como outros modelos de funções demográficas, incorpora seletiva por sexo e idade (Amaral, et.al, 2004). A seletividade por sexo é ajustada no modelo ao incorporar no modelo feminino um pico mais precoce de migração, que segundo os autores captaria a diferença na idade do casamento. A seletividade de idade é muito mais bem definida, com o pico de migração acontecendo nas idades mais jovens. Os jovens adultos, entre 20 e 25 anos, apresentam as maiores taxas de migração, enquanto os jovens adolescentes, entre 10 e 19 anos, apresentam as menores taxas. As taxas de migração das crianças (0 a 9 anos) são um reflexo do fluxo migratório de seus pais (26 a 40 anos). Os padrões que tratam de movimentos originados ou destinados a grandes cidades, com altos níveis de serviços e comodidades culturais, apresentam ainda um pico próximo à idade de 65 anos.

Nesse sentido, o padrão de migração proposto (Rogers e Castro 1981) é composto de quatro componentes relacionados ao mercado de trabalho: (1) a primeira parte da curva antes da entrada no mercado de trabalho é uma curva exponencial negativa de zero a 19 anos-de-idade (α_1 como o indicador descendente); (2) a curva para os migrantes em idade de trabalho tem uma forma parabólica (μ_2 como o indicador de média idade; λ_2 como o indicador ascendente e α_2 como o indicador descendente); (3) a curva de pós-trabalho (aposentadoria) é uma pequena parábola significando as pessoas em torno de 65 anos de Idade (μ_3 como o indicador de média de idade; λ_3 como o indicador ascendente e α_3 como o indicador descendente); e (4) o último

parâmetro da programação modelo é uma constante (c), que ajusta as taxas de migração para a expressão matemática. Essa proposição estabelece que a migração é altamente influenciado pela economia, porque as curvas designar diferentes momentos de entrada de um indivíduo no mercado de trabalho.

O modelo de perfil etário da migração é composto por três funções diferentes usadas para modelar as taxas de migração (Rogers e Castro, 1981). O modelo básico de migração tem todos os quatro componentes de migração laboral listados acima. Um segundo modelo, conhecido como forma reduzida, apresenta a curva de aposentadoria para os indivíduos em torno de 65 anos de idade. Um terceiro modelo, migração com inclinação ascendente, tem uma curva linear para idades acima da idade ativa, ao invés de uma parábola. Os fluxos migratórios são geralmente modelados utilizando o "modelo reduzido", porque não há nenhuma função parábola ou linear evidente para as idades após a saída dos trabalhadores do mercado de trabalho (Raymer e Rogers de 2007; Rogers e Jordan, 2004).

Rogers e Castro (1984) realizaram também uma análise comparativa entre padrões migratórios nacionais e internacionais, e identificaram três famílias distintas de modelos-padrão por idade. A primeira família foi chamada de “modelo-padrão básico de migração”, que ilustra exatamente a curva descrita anteriormente, com a soma de quatro componentes, totalizando 11 parâmetros. O segundo grupo é a “forma reduzida”, que agrega as funções que não apresentam a parábola nas curvas pós- laborais, com um total de 7 parâmetros. Por fim, o terceiro perfil é o “modelo-padrão de migração com uma inclinação ascendente”, que soma 9 parâmetros. Cada família de modelo de migração disponibiliza parâmetros que permitem classificar diferentes categorias de perfis migratórios. A escolha da família-padrão para a modelagem deve ser feita de acordo com o comportamento das taxas de migração observadas de cada região de origem para cada região de destino.

A Figura 1 ilustra a curva típica da primeira família de modelos de migração e também mostra um modelo de Rogers-Castro adaptado com um termo exponencial dupla para a migração de retorno centrada em 55 anos.

Figura 1. Curva do modelo migratório de Rogers e Castro

Fonte: Hill e Queiroz (2010).

Um modelo mais recente, proposto por Raymer e Rogers (2007), propõem a estimação das estruturas por idade e espaciais de migração inter-regional com estimativas indiretas. O intuito é de propiciar técnicas para estimação de migração para aqueles que utilizam dados de baixa qualidade. Além disso, países podem apresentar a perda de informações detalhadas de migração, como é o caso dos Estados Unidos que não terão mais informações sobre fluxos populacionais nos dados censitários do questionário ampliado, o qual será substituído por uma pesquisa mensal contínua (“American Community Survey”). Como a migração possui fortes regularidades espaciais e por idade, um modelo multiplicativo log-linear é proposto. O componente multiplicativo pode ser dividido em:

- Efeito geral (“overall”): componente que representa o nível da migração.
- Origem: componente que representa as “repulsões” relativas de cada região.
- Destino: componente que representa as “atrações” relativas de cada região.
- Interação origem-destino: componente que representa os impactos de distância física ou social entre lugares (e não explicados pelos efeitos acima).
- Informação de idade do migrante também pode ser utilizada nessa modelagem.

Este componente é estimado a partir do seguinte modelo:

$$n_{ijx} = (T)(O_i)(D_j)(A_x)(OD_{ij})(OA_{ix})(DA_{jx})(ODA_{ijx}), \quad (1)$$

onde n_{ijx} : migração observada da região i para a região j na idade x ; T : numero total de migrantes (n_{++}); O_i : proporção de todos os migrantes deixando a região i (n_{i+}/n_{++}); D_j : proporção de todos os migrantes mudando para a região j (n_{+j}/n_{++}); OD_{ij} : razão entre migração observada e migração esperada $n_{ij}/[(T)(O_i)(D_j)]$; e A_x : proporção de todos os migrantes no grupo etário x .

O modelo log-linear saturado a ser estimado tem a seguinte forma:

$$\ln(n_{ijx}) = \lambda + \lambda_i^O + \lambda_j^D + \lambda_x^A + \lambda_{ij}^{OD} + \lambda_{ix}^{OA} + \lambda_{jx}^{DA} + \lambda_{ijx}^{ODA} \quad (2)$$

3. Uma aplicação de Rogers e Castro para o Brasil

Uma das importantes aplicações do modelo de migração proposto por Rogers e Castro é para ajustar a função de migração quando os dados são de baixa qualidade ou quando está-se usando dados de pequenas áreas que podem apresentar grandes flutuações. Amaral, Figoli e Nascimento (2004), Amaral (2008) e Amaral, Rios-Neto e Potter (2011) apresentam diversas estimativas das funções de migração para o Brasil a partir do modelo matemático. O trabalho de Amaral, Rios-Neto e Potter (2011) é um estudo de caso interessante pois estima as funções de migrações para 502 microrregiões brasileiras em diferentes pontos do tempo.

No trabalho, a metodologia envolveu, primeiramente, gerar uma estimativa de migração a partir dos dados censitários brasileiros. No caso, os autores utilizaram a pergunta sobre migração de data-fixa, residência 5 anos antes do censo, presente nos censos demográficos de 1991 e 2000. A abordagem neste estudo requer a estimativa de imigração, em vez de emigração como um método para obter o impacto de fluxos populacionais nas variações populacionais no final do período (Amaral, Rios-Neto e Potter, 2011). Assim, a taxa específica para a idade da migração interna por faixa etária foi estimada entre as cinco grandes regiões brasileiras em dois Censos.

$$Taxa\ Imigração\ por\ idade\ (TMI_{x,ij}) = \frac{\Sigma(K_{ij}^x)}{t * \Sigma \left[\frac{(K_{j.}^x + K_{jj}^x) + (K_j^x)}{2} \right]} \quad (3)$$

Segundo os autores, foi necessário calcular a TMI proporcional para gerar o padrão de migração por faixa etária, assim, a soma das taxas é igual a 1, quando se considera a região de origem, de destino e o ano censitário (Amaral, Rios-Neto e Potter, 2011). No trabalho, centrado

em entender a dinâmica demográfica do mercado de trabalho brasileiro, os autores geral estimativas de migração para 10 faixas etárias (15-19, 20-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50-54, 55-59, e 60-64).

Os autores usaram a modelagem de taxas de migração baseada no segundo componente (migrantes em idade de trabalho) e quarto componentes (constante para ajustar a expressão matemática) do modelo de Rogers e Castro. Depois da estimação das taxas de imigração por grupo etário, o modelo matemático proposto por Rogers e Castro (1981) foi implementado com o Table Curve 2D. Há o módulo de regressão não-linear do SPSS, com o método de estimação Levenberg-Marquardt. É possível também fazer essa modelagem com o TableCurve 2D e o Matlab. Ou seja, foi utilizado o modelo reduzido apresentado anteriormente, seguindo a equação:

$$S_x = a_2 * \exp\{-\alpha_2(x-\mu_2) - \exp[-\lambda_2(x-\mu_2)]\} + c. \quad (4)$$

As taxas proporcionais (ASIMRx, ij) por faixa etária, migração foram estimados com a Equação (1) e modelado com a equação (2), para os fluxos de população entre regiões brasileiras. A Figura 2 ilustra o perfil etário da migração gerado pelo modelo proposto do Nordeste para o Sudeste e a Figura 3 apresenta as estimativas do perfil etário do Norte para o Sudeste. Os pontos são as taxas calculadas com os dados diretos, e a linha representa o ajuste do modelo. No caso do fluxo Nordeste-Sudeste, os dados observados são bastante próximos aos ajustados pelo modelo. No caso do fluxo Norte-Sudeste o modelo de Rogers e Castro permite uma significativa suavização das taxas observadas, chegando a resultados mais robustos. Os resultados também indicam a importância do método para estudos de migração em áreas com dados deficitários ou com um pequeno número de observações.

Figura 2. Perfil etário da bigração, Nordeste para Sudeste, observado e ajustado, 1991 e 2000

Fonte: Amaral, Rios-Neto e Potter (2011).

Figura 3. Perfil etário da bigração, Norte para Sudeste, observado e ajustado, 1991 e 2000

Fonte: Amaral, Rios-Neto e Potter (2011).

4. Aplicando Rogers e Castro para o estudo de mortalidade

Para avaliar a cobertura de mortes relatadas usamos tradicionais métodos demográficos, chamada Distribuição Morte Métodos-DDM (Hill, You e Choi, 2009; Dorrington, 2012a; 2012b) De modo geral, há três principais métodos de avaliação de registro de óbitos: *general growth balance* (GGB) proposto por HILL (1987), *synthetic extinct generation* (SEG) proposto por Bennett e Horiuchi (1981) e o *adjusted synthetic extinct generations* (SEG-adj) proposto por Hill, You e Choi (2009). Os DDM são comumente usados para estimar a mortalidade de adultos em uma população não-estável (Timeaus, 1991; Hill et al, 2005; Hill, et.al, 2009) e fazem diversas premissas fortes: 1) que a população é migração ou observe fluxos bem pequenos, 2) qualidade da declaração dos óbitos por idade é constante; 3) alidade da declaração da população por idade é constante e 4) idades dos vivos e dos óbitos sejam relatadas sem erros.

O método General Growth Balance é estimado a partir da seguinte equação:

$$b(a+) - r^o(a+) = (1/t) \ln(k) + qd * d^o(a+), \quad (5)$$

onde $b(a+)$ e $a(+)$ são a taxa de entrada e saída na população, $r^o(a+)$ é a taxa de crescimento observada, t é o intervalo de intercensal em anos, k é a cobertura do primeiro censo relação ao segundo e qd denota o grau de cobertura da notificação de óbitos em relação à cobertura média da população. A estimativa da cobertura de morte, e fator de ajuste, é dada pela inclinação da linha referente as taxas de mortalidade estimadas com estrutura etária da população e as taxas de mortalidade observadas. É importante observar que no método proposto assume-se que população fechada.

A suposição de que a população está fechado para a migração é muito forte quando se está trabalhando com pequenas áreas ou lidando com fluxos migratórios internacionais. Uma alternativa é assumir que existe alguma x idade acima que a migração líquida é desprezível. Assim, o desempenho dos métodos acima que a idade não será afetada (Hill, You e Choi, 2009; Murray, et.al, 2009; Dorrington, 2012).

Uma segunda alternativa, proposta por Bhat (2002) é incluir as taxas de migração nas estimativas das taxas de mortalidade residuais do modelo, conforme apresentado nesta equação:

$$r(a+) = b(a+) - d(a+) + nm(a+), \quad (6)$$

onde $nm(a+)$ é a taxa líquida de migração para a população com idade a e superior. A maior limitação de utilizar a metodologia proposta por Bhat (2002) é a necessidade de ter informações sobre a migração por idade. Com base nisso, Hill e Queiroz (2010) propuseram uma modificação ao método General Growth Balance que incorpora as taxas de migração a partir do modelo proposto por Rogers e Castro (1994).

A proposta de Hill e Queiroz (2010) faz a seguinte modificação na equação básica do modelo General Growth Balance. A nova equação assume que as taxas de migração têm um perfil etário bem definido nm^s_x . Dessa forma, os autores reescrevem a Equação (6) da seguinte forma:

$$r(a+) + d(a+) - b(a+) = k_{nm} + q_{nm} * nm^s(a+), \quad (7)$$

onde:

$$nm^s(a+) = \sum nm^s_x * {}_5PYL_x / \sum {}_5PYL_x. \quad (8)$$

Os autores definem q_{nm} como uma constante que relaciona a migração do modelo padrão com o que seria observado na população. Hill e Queiroz (2010) mostram que a inclinação da reta apresentada na Equação 7 é uma estimativa de q_{nm} . Com esses valores, eles propõe uma nova equação para estimar o modelo e assim o grau de cobertura de mortalidade.

$$b(a+) - r^o(a+) + (k_{nm} + q_{nm} * nm^s(a+)) = (1/t) \ln(k) + q_d * d^o(a+) \quad (9)$$

Como forma de generalizar o método proposto, Hill e Queiroz (2010) sugerem usar o modelo de Rogers e Castro (1994) para gerar uma estimativa do padrão de migração que permite um bom ajuste do modelo de correção de sub-registro. O modelo, no entanto, é um de um fluxo de migração bruto. O padrão etário da migração líquida provavelmente se assemelham a de um fluxo bruto somente se o fluxo é muito grande em uma direção.

Com base no modelo e parâmetros básicos sugeridos por Rogers e Castro (1984), o modelo simplificado de migração pode ser estimado da seguinte forma (Equação 10). Os autores também aplicam o modelo modificado de Rogers e Castro para levar em consideração a migração de retorno nas idades mais avançadas (Equação 11).

$$m(x) = 0,02 * \exp(-0,1 * x) + 0,1 * \exp(-0,1 * (x-20) - \exp(-0,4 * (x-20))) \quad (10)$$

$$m(x) = 0.02 * \exp(-0.1 * x) + 0.1 * \exp(-0.1 * (x-20) - \exp(-0.4 * (x-20))) - 0.02 * \exp(-0.125 * (x-55) - \exp(-0.4 * (x-55))) \quad (11)$$

Assim, Hill e Queiroz (2010) sugerem que as taxas específicas sejam calculadas para grupos etários quinquenais, como no exemplo de Amaral, et.al (2011) e que a Equação (10) seja avaliada para as faixas etárias de 5 anos. As estimativas obtidas, segundo os autores, fornecem taxas de migração líquida e, assim, podem ser incorporadas no modelo GGB (Hill e Queiroz, 2010). Bell (2005) argumenta que o modelo de Rogers e Castro não propicia parâmetros adequados, já que os mesmos podem apresentar valores semelhantes para curvas de migração distintas e vice-versa. Segundo ele, os parâmetros não conseguem realmente descrever o padrão e nível das curvas de migração, como ocorre com parâmetros de modelos para fecundidade e mortalidade. Nesse caso, o modelo funciona melhor para suavizar curvas ou ajustar dados deficientes. Todavia, a abordagem apresentada por Hill e Queiroz (2010) apresentou resultados satisfatórios.

Hill e Queiroz (2010) apresentam resultados dos fatores de cobertura dos registros vitais para México e Porto Rico. A Figura 4 mostra os resultados do General Growth Balance para o México. A variação em relação ao modelo de ajuste nas idades iniciais indica possíveis efeitos de fluxos migratórios nos métodos de correção de sub-registro (Hill, You e Choi, 2009; Hill e Queiroz, 2010).

Figura 4. Gráfico de ajuste do modelo “general growth balance”, homens, México, 1990-2000

Fonte: Adaptado e retirado de Hill e Queiroz (2010).

A Figura 5 de Hill e Queiroz (2010) apresenta a intensidade da migração por idade $x+$ estimada a partir do modelo proposto por Rogers e Castro. O eixo vertical apresenta a estimativa residual da taxa de migração e no eixo horizontal a taxa de migração derivada do modelo de Rogers – Castro, ajustado a distribuição idade apropriada. O passo seguinte, na metodologia proposta por Hill e Queiroz (2010) é aplicar os valores estimados de $(k_{nm} + q_{nm} * n_{ms}(a+))$ para ajustar a estimativa residual lado esquerdo da taxa de mortalidade das idades $a+$ para os efeitos de migração. A Figura 6 de Hill e Queiroz (2010) apresenta as reestimativas do General Growth Balance ajustando a Equação 6 às taxas de migração obtidas após aplicar o modelo de Rogers e Castro. A análise gráfica indica que o modelo o ajuste global do GGB melhorou substancialmente. De particular interesse é a observação de que o intercepto do modelo explica razoavelmente bem a migração de retorno.

Figura 5. Gráfico da intensidade da migração, ajustado pelo modelo de Rogers e Castro, homens, México, 1990-2000

Fonte: Adaptado e retirado de Hill e Queiroz (2010).

Figura 6. Gráfico de ajuste do modelo “general growth balance”, ajustado por migração, homens, México, 1990-2000

Fonte: Adaptado e retirado de Hill e Queiroz (2010).

5. Modelando o padrão etário da migração recente – Xu-Doeve

Xu-Doeve (2005, 2006, 2007) argumenta que os recentemente se instalar em uma área são muitas vezes perdida pelos censos demográficos. E que é possível estimar esse grupo migratório, retroprojetando a população recenseada e ajustar um modelo em uma função quadrática ou cúbica e derivando os números de migração a partir desta curva ajustada.

Xu-Doeve (2005, 2006, 2007) argumenta que definindo $K_0(t)$ como a população de uma área que ainda não experimentou migração no tempo t antes do censo, então

$K_0(t) = K_0(0) \exp \left[- \int_0^t \mu(u) du \right]$ onde $\mu(u)$ é a força da imigração no tempo u antes do censo.

Assim, $\ln(K_0(t)) = \ln(K_0(0)) + \left[- \int_0^t \mu(u) du \right]$. $K_0(0)$ é a população da área na data do censo e $K_0(t)$ pode ser obtida por $K_0(t-1)$ ao subtrair o número de imigrantes entre $t-1$ e t indo para anos passados.

Xu-Doeve (2005) também argumenta que $\mu(u)$ pode assumir a forma quadrática, cúbica ou um polinômio de quarto grau que pode ser ajustado pelo modelo de regressão na variável $\ln(K_0(t))$ em t, t^2, t^3 , etc. Além disso, se houver indicativo que $K_0(0)$ foi subestimado devido a perda de imigrantes no censo demográfico, é possível limitar o modelo de regressão a alguns

pontos anteriores ao censo e estimar o modelo considerando apenas esses dados. A estimativa do modelo irá prover uma estimativa de $K_0(0)$ corrigida pelos migrantes perdidos.

O modelo proposto é derivado da seguinte forma. Seja $K_0(t)$ o número pessoas que não migrou no tempo t :

$$K_0(t) = K_0(0) \exp \left\{ \int_0^t -(\mu_m(t) + \mu_d(t)) dt \right\}, \quad (12)$$

onde $K_0(t)$ é o tamanho da coorte e $\mu_{\cdot B}(t)$ é o função de risco do evento ocorrer que mede a intensidade do evento e é instantânea. Com essas informações podemos acompanhar a mudança do tamanho de uma coorte. A analogia mais direta que podemos fazer é com a evolução do tamanho de uma coorte que vai se reduzindo devido aos efeitos da força de mortalidade em cada idade.

No caso acima, a estimativa de pode ser retroprojetada, assim o movimento de imigração pode ser transformado em emigração. Para estimar o modelo proposto por Xu-Doeve é necessário ter informações sobre a história de vida dos membros de uma coorte até, pelo menos, o primeiro evento. No caso de migração, em diversos países os censos demográficos incluem uma pergunta sobre tempo de residência na localidade, e essa informação pode ser usada para fazer a estimativa proposta. Substituindo por B e ajustando a população pela função de sobrevivência teremos: $K_0(t) = K_0(0) \exp \left\{ \int_0^t -\mu_{\cdot B}(t) dt \right\}$. Assumindo que $\mu_{\cdot B}(t) = \beta_1 + \beta_2 t + \beta_3 t^2 + \dots$.

Em palavras, para gerar as estimativas do modelo proposto por Xu-Doeve (2005) é necessário organizar os dados por coorte e duração de residência no local de interesse. Como o perfil de migração pode ser bastante distinto por sexo, o autor sugere que as tabulações sejam apresentadas separadamente para homens e mulheres. Se possível, além da informação de duração de residência atual, poderia ser coletada a informação de residência anterior.

No modelo é importante considerar, como exposto, que os indivíduos estão sujeitos a dois riscos: migração e morte. Dessa forma, os dados precisam ser ajustados pelas probabilidades de sobrevivência. A forma mais direta é aplicar um modelo de retroprojeção com base em uma tabela de vida adequada. Em outras palavras, a ideia é reconstruir uma coorte antes que ela tenha sido exposta ao risco de mortalidade.

Após essa correção, os dados apresentados devem ser acumulados do tempo zero (0) até o tempo mais distante de residência. A organização dos dados dessa maneira, informa as mudanças no tamanho das coortes como consequência da imigração a medida que o tempo passa. Entretanto, precisamos reverter o tempo para que a organização dos dados reflita o arcabouço teórico apresentado nas equações iniciais. Ou seja, o último ano de dados se transforma no primeiro ponto de observação e assim sucessivamente. Por exemplo, se temos a informação construída a partir do censo de 2010. O ano de 2010 se transforma em $t = 0$, 2009 em $t=1$ e assim sucessivamente.

Temos que $\ln(K_0(t)) = \beta_0 + \beta_1(-t) + \beta_2(-\frac{1}{2}t^2) + \beta_3(-\frac{1}{3}t^3)$. Dessa forma podemos ajustar, $\ln(K_0(t))$, para todo valor de $t > 0$ e excluindo o valor de K quando $t = 0$. Os dados organizados da forma descrita anteriormente permitem a construção das informações para cada coorte em cada ponto do tempo definido. Tirando o logaritmo de cada uma dessas informações obtém-se a variável resposta da regressão acima. O ajuste pode ser feito com um modelo de regressão linear básico. Os resultados da regressão permitem recuperar o tamanho inicial de cada uma das coortes, a partir das estimativas de β_0 , as estimativas da intensidade da migração β_1 , e o que vem acontecendo com o fluxo migratório ao longo do tempo β_2 .

A Tabela 1 foi construída com base no Censo da África do Sul de 1996 e contém a informação do número de imigrantes por idade em relação ao ano de migração.

Tabela 1. Número de imigrantes depois de ajustado pela função de mortalidade

Número de Imigrantes após corrigir pela função de mortalidade							
Ano Calendário da Mudança							
Idade	1996	1997	1998	1999	2000	2001	Total
5-9	2,401	10,899	12,014	13,899	15,699	16,112	70,823
10-14	2,730	10,514	12,427	15,013	15,119	15,597	71,333
15-19	2,687	10,859	13,530	16,734	18,980	27,547	90,151
20-24	3,114	10,936	15,667	20,235	25,693	40,913	116,016
25-29	3,311	13,457	17,881	20,092	27,975	37,105	118,589
30-34	3,341	11,397	14,743	16,580	19,573	25,792	89,852
35-39	3,326	10,170	12,458	13,138	15,908	18,018	71,335
40-44	2,188	7,479	7,937	9,495	10,630	13,385	49,856
45-49	1,554	5,172	5,673	6,853	6,756	8,127	33,237
50-54	1,122	3,119	3,859	4,827	4,915	6,147	23,305
55-59	811	2,234	2,868	3,673	3,240	3,827	16,041
60-64	665	2,036	2,421	3,003	3,116	4,117	14,634
65-69	542	1,900	1,535	2,240	2,119	2,388	9,959
70-74	474	1,042	1,327	1,308	1,836	1,940	7,178
75-79	337	834	710	1,017	1,141	1,492	4,806
80-84	218	816	493	715	714	752	2,961
85+	317	556	474	271	261	606	1,681
Total	29,139	103,421	126,019	149,092	173,676	223,866	791,756

A Tabela 2 apresenta os valores acumulados de imigrantes construídos a partir da informação obtida na Tabela 1. A Tabela 2 informa as mudanças no tamanho das coortes como consequência da imigração a medida que o tempo passa.

Tabela 2. Número acumulado de imigrantes

Idade	2001	2000	1999	1998	1997	1996	Total
	0.772603	1.772603	2.772603	3.772603	4.772603	5.772603	
0-4							
5-9	1,098,595	1,082,895	1,068,996	1,056,982	1,046,083	1,043,683	1,114,707
10-14	1,121,493	1,106,374	1,091,361	1,078,934	1,068,419	1,065,689	1,137,090
15-19	1,098,295	1,079,314	1,062,581	1,049,050	1,038,191	1,035,504	1,125,842
20-24	865,631	839,938	819,703	804,036	793,100	789,986	906,544
25-29	752,587	724,612	704,520	686,639	673,182	669,871	789,692
30-34	607,235	587,662	571,082	556,339	544,942	541,601	633,027
35-39	568,150	552,242	539,105	526,646	516,476	513,150	586,168
40-44	478,030	467,399	457,905	449,968	442,489	440,300	491,414
45-49	391,435	384,679	377,825	372,152	366,980	365,426	399,562
50-54	325,900	320,985	316,158	312,299	309,180	308,058	332,047
55-59	236,626	233,386	229,713	226,845	224,612	223,800	240,454
60-64	213,193	210,076	207,073	204,652	202,616	201,951	217,309
65-69	151,824	149,705	147,465	145,930	144,030	143,488	154,213
70-74	130,750	128,914	127,605	126,278	125,236	124,762	132,690
75-79	70,241	69,100	68,083	67,373	66,539	66,202	71,732
80-84	49,663	48,949	48,234	47,741	46,925	46,707	50,415
85+	27,353	27,092	26,821	26,347	25,791	25,474	27,959

O modelo ajustado por meio de regressão simples é apresentado na Tabela 3. O valor do exponencial de β_0 é uma estimativa do tamanho inicial da população e permite ajustar o valor caso o número inicial de imigrantes tiver sido sub-estimado. O valor de β_2 negativo para quase todas os grupos de idade indica que voltando no tempo a força de migração é menor. Ou seja,

com a passage do tempo o ritmo de migração em direção a localidade aumenta. A função hazard é calculada pela combinação das estimativas dos β .

Tabela 3. Ajuste do modelo

β_3	β_2	β_1	$\ln(K_0(0))$	R2
0.00016	-0.00223	0.01704	13.92208	0.999994
-0.00025	0.00009	0.01420	13.94120	0.999837
0.00000	-0.00234	0.02061	13.92454	0.999935
0.00038	-0.00735	0.03897	13.69922	0.999882
0.00105	-0.01130	0.04975	13.56636	0.999832
0.00013	-0.00424	0.03790	13.34471	0.999845
0.00060	-0.00568	0.03443	13.27512	0.999734
0.00063	-0.00537	0.02869	13.09819	0.99985
-0.00005	-0.00087	0.01920	12.89228	0.999672
-0.00009	-0.00109	0.01728	12.70748	0.999417
-0.00065	0.00232	0.01270	12.38481	0.999071
-0.00001	-0.00147	0.01706	12.28279	0.999609
0.00066	-0.00445	0.01967	11.94468	0.998866
0.00084	-0.00604	0.02018	11.79493	0.999036
0.00141	-0.00927	0.02675	11.17801	0.998866
0.00178	-0.00964	0.02494	10.82992	0.998698
0.00217	-0.00674	0.01452	10.22616	0.99648

Uma análise importante do modelo é verificar se ele permite obter uma estimativa adequada do tamanho inicial, ou seja, a estimativa do modelo irá prover uma estimativa de $K_0(0)$ corrigida pelos migrantes perdidos. A Figura 7 mostra a comparação entre os valores observados do grupo de 20 a 25 anos de idade e a estimativa gerada pelo modelo. Para o grupos etário 20–25 o volume da migração usando é subestimado (menor do que o observado), sugerindo que neste caso o método pode não ter corrigido adequadamente a subestimação de migrantes como sugerido. Entretanto, outros exemplos apresentados pelo autor (Xu-Doeve, 2008) mostram resultados bastante robustos e na direção do esperado pelo modelo proposto.

Figura 7. Comparação do valor observado e estimado do grupo etário de 20 a 25 anos

Nessa mesma linha, um exemplo interessante pode ser visto nos trabalhos de Schmertmann (1992; 1999), que não serão discutidos em detalhes nesse capítulo. Schmertmann (1992) usa informações sobre duração da residência para estimar as taxas de migração a partir de um único censo. O método proposto não substituiu estimativas feitas com múltiplos censos, mas oferece uma alternativa para o estudo do processo migratório quando se tem limitações de dados demográficos. O método produz estimativas de taxas de migração comparando informações de taxas de migração ao longo da vida com estimativas de tendências temporais de diferentes taxas de migração.

Schmertmann (1999) usa a informação de migração de última etapa e propõe um método para estimar transições a partir dessas informações. O modelo usa retroprojeções para produzir taxas de migrações e a aplicação, com base em dados do Brasil, mostrou-se bastante robusta.

6. Conclusão

Este capítulo buscou discutir as técnicas usadas para estimar as transições de migração a partir de diferentes modelos matemáticos, apresentar exemplos de suas aplicações para alguns países e outras aplicações em estudos demográficos (mortalidade). O esclarecimento das limitações anteriores sobre estimativas de migração, bem como a elucidação de melhorias necessárias para estimar as taxas são importantes, pois oferecem a diferentes pesquisadores mecanismos e procedimentos adequados que podem ser usados com outros censos demográficos e pesquisas, mesmo que limitadas, que contém informação de migração.

Os modelos apresentados nessa capítulo são um recurso para verificar como estimar as taxas de migração com base em dados limitados (Bell 2005), permitindo o estudo dos níveis e padrões de migração em vários países. Os exemplos mencionados, também, permitem calcular os fluxos migratórios com base em informações sobre os movimentos de última etapa, data fixa e duração da residência. A análise desses fluxos é uma ferramenta importante para entender melhor os níveis de migração entre diferentes áreas. As múltiplas possibilidades de medição migração oferecida por informações sobre dados de última movimentação e duração da residência (Xu-Doeve 2006) foram explorados e desenvolvidos por demógrafos ao longo do tempo. Desenvolvimentos adicionais em metodologias para estimar as taxas de migração, bem como a criação de novas técnicas são possíveis através da utilização de dados reais e da aplicação de tais procedimentos em contextos específicos (Rogers e Castro, 1981; Schmertmann 1999).

Estudos adicionais poderiam melhorar a qualidade de técnicas e resultados obtidos pela investigação nesta área, permitindo que os demógrafos para entender melhor os níveis de migração e padrões. Tais melhorias futuras em estudos de migração ajudaria gerar realizações científicas adicionais, semelhantes aos benefícios obtidos pelos estudos de padrões de mortalidade e fecundidade.

7. Referências

- Amaral, E. F. L. 2008. "Improvements of techniques to estimate migration rates: an application with Brazilian censuses data." *Population Review*, 47(2): 1-24.
- Amaral, E.F.L, Rios-Neto, E.L.G e Potter, J. 2011. "Influences of transition in age-education structure and internal migration on the labor market in Brazil." *Revista. gepec* 85903: 000.
- Amaral, E. F., do Nascimento Rodrigues, R., & Fígoli, M. G. B. (2013). Métodos e técnicas de mensuração de movimentos migratórios: Goiás e Distrito Federal, 1975-1979 e 1986-1990. *Revista Brasileira de Estudos de População*, 21(2), 283-301.
- Bhat, M. General growth balance method: a reformulation for populations open to migration. *Population Studies*, 56(1), 23-34, 2002.
- Bell, M., Blake, M., Boyle, P., Duke-Williams, O., Rees, P., Stillwell, J. 2002. "Cross-national comparison of internal migration: issues and measures." *Journals of the Royal Statistical Society*, Series A, 165(3): 435–464

- (http://espace.library.uq.edu.au/eserv/UQ:62132/Bell_2002_Crossnational_comparisons.pdf).
- Bell, Martin. 2005. "Towards Rigorous Cross-National Comparison of Internal Migration: Who Collects What?" Paper presented at the XXV International Population Conference of the International Union for the Scientific Study of Population (IUSSP), Tours, France, July 18 to 23, 2005.
- Beltrão, K. I. e M. H. F. T. Henriques. 1987. "Modelagem da migração líquida rural-urbana no Brasil: décadas de 1960/1970 e 1970/1980." *Previdência em Dados*, 2, 3: 23–36.
- Bennett, N.; Horiuchi, S. Estimating the completeness of deaths registration in a closed population. *Population Index* 47(2): 202-221, 1981.
- Bracken, I., Bates, J. 1983. "Analysis of gross migration profiles in England and Wales: some developments in classification." *Environment and Planning A*, 15(3): 343–355 (<http://www.envplan.com/abstract.cgi?id=a150343>).
- Hill, K.. Estimating census and death registration completeness. *Asian and Pacific Population Forum*, 1(3): 8-12, 1987.
- Hill, K., & Queiroz, B. (2010). Adjusting the general growth balance method for migration. *Revista Brasileira de Estudos de População*, 27(1), 7-20.
- Hill, K; You, D.; Choi, Y. Death distribution methods for estimating adult mortality: sensitivity analysis with simulated data errors. *Demographic Research*, v. 21(9): 235-254, 2009a.
- Hill, K.; Choi, Y.; Timaeus, I. Unconventional approaches to mortality estimation. *Demographic Research*, v. 13(12):281-300, Nov. 2005.
- Machado, C. C. 1993. "Projeções multirregionais de população: o caso brasileiro (1980–2020)." Belo Horizonte: Centro de Desenvolvimento e Planejamento Regional (CEDEPLAR), Universidade Federal de Minas Gerais (UFMG). (Tese de doutorado)
- Machado, C. C. 2005. "Indirect estimation of migration flows: application of new approaches to Brazilian census data 1991-2000." Apresentado na XXV Conferência de População Internacional da União Internacional para o Estudo Científico da População (IUSSP), Tours, França.
- Raymer, J., A. Bonaguidi, e A. Valentini. 2006. "Describing and projecting the age and spatial structures of interregional migration in Italy." *Population, Space and Place*, 12, 5: 371–388.
- Raymer, J., e A. Rogers. 2007. "The American community survey's interstate migration data: strategies for smoothing irregular age patterns." Apresentado no encontro anual da Associação de População da América (PAA), New York, NY, Estados Unidos.

- Raymer, J., e A. Rogers. 2007. "Using age and spatial flow structures in the indirect estimation of migration streams." *Demography*, 44, 2: 199–223.
- Rogers A, F. J. Willekens, e J. Raymer. 2003. "Imposing age and spatial structures on inadequate migration-flow datasets." *The Professional Geographer*, 55, 1: 56–69.
- Rogers, A., & Castro, L. J. (1984). Model migration schedules. *Migration, urbanization, and spatial population dynamics*, 41-91.
- Rogers, A., e L. J. Castro. 1981. Model migration schedules. Laxenburg, Áustria: International Institute for Applied Systems Analysis.
- Rogers, A., e L. Jordan. 2004. "Estimating migration flows from birthplace-specific population stocks of infants." *Geographical Analysis*, 36, 1: 38–53.
- Rogers, A., F. J. Willekens, e J. Raymer. 2001. "Modeling interregional migration flows: continuity and change." *Mathematical Population Studies*, 9, 3/4: 231–263.
- Rogers, A., F. J. Willekens, J. S. Little, e J. Raymer. 2002. "Describing migration spatial structure." *Papers in Regional Science*, 81, 1: 29–48.
- Schmertmann, Carl P. (1999). "Estimating Multistate Transition Hazards From Last-Move Data." *Journal of the American Statistical Association* 94 (445): 53–63.
- Schmertmann, C. P. (1992). Estimation of historical migration rates from a single census: Interregional migration in Brazil 1900–1980. *Population Studies*, 46(1), 103-120.
- Xu-Doeve, William L. J. 2005. "The demographic measurement of migration and its adjustment for underenumeration." Paper presented at the XXV International Population Conference of the International Union for the Scientific Study of Population (IUSSP), Tours, France, July 18 to 23, 2005.
- Xu-Doeve, William L. J. 2006. "The Measurement of International and Internal Migration in the 2010 Global Round of Population Censuses: Twelve Key Recommendations on Questions, Concepts and Procedures." Paper presented at the European Population Conference of the European Association for Population Studies (EAPS), Liverpool, United Kingdom, June 21 to 24, 2006.
- Xu-Doeve, William L. J. 2008. "Introduction to the measurement of internal and international migration." Netherlands: ANRC Publishing.