

# A TÁBUA DE VIDA

### 9.1 INTRODUÇÃO

O capítulo anterior introduziu vários indicadores para quantificar a mortalidade, começando pela Taxa Bruta de Mortalidade (TBM) e passando por vários tipos de Taxas Específicas por idade e sexo. Entretanto, os indicadores introduzidos até agora, que quantificam o risco de morte em diferentes idades, não respondem todas as perguntas relevantes para caracterizar a mortalidade. Por exemplo, qual é a probabilidade de que uma pessoa da idade  $x$  que está viva hoje ainda estará viva em 20 anos? Ou, mais especificamente, de todas as pessoas que nascem, quantas chegarão à idade de 65 anos? Ou, prevalecendo as condições atuais de mortalidade, qual é o número esperado de óbitos de mulheres entre 15-49 anos de idade, de uma coorte hipotética de 100.000 nascimentos? Ou quantos anos as pessoas vivem em média? Para muitas aplicações práticas da análise demográfica são estas as perguntas que importam e uma mera sequência de Taxas Específicas de Mortalidade (TEMs) não é suficiente para respondê-las. É preciso dispor de um instrumento mais abrangente que quantifica diversos aspectos ou diversas *funções* da mortalidade que interessam do ponto de vista das principais aplicações. Este instrumento é conhecido como a *tábua de vida*, também chamada *tabela de vida*, *tábua de mortalidade* ou *tabela de sobrevivência*. Trata-se de um instrumento antigo. Como já foi observado no Capítulo 5, considera-se que a primeira tábua de vida, da Cidade de Londres, foi publicada por John Graunt, em 1662. No seu livro *Natural and Political Observations Made upon the Bills of Mortality*, Graunt apresentou o cálculo de uma coorte inicial de 100 pessoas, de que 64 chegaram à idade de 6 anos, 40 à idade de 16 e apenas 25 à idade de 26 anos exatos. Em 1693, Edmond Halley aperfeiçoou a técnica, com base nas estatísticas

da Cidade de Breslau, na Alemanha, e apresentou um mecanismo de cálculo muito parecido com a forma como se calcula a tábua de vida hoje em dia.

As tábuas de vida podem ser *completas* (para idades 0, 1, 2, 3, 4, 5 etc.) ou *abreviadas* (para idades 0, 1, 5, 10, 15 etc.). As tábuas abreviadas são mais comuns. Por um lado, isso é porque uma tábua completa exige muita informação cuja qualidade precisa ser muito boa; as exigências de dados para as tábuas abreviadas são menores. Por outro lado, as tábuas completas são muito grandes e fornecem mais informação do que a maioria dos usuários precisa. Devido às diferenças consideráveis que costumam existir entre a mortalidade masculina e feminina, as tábuas de vida normalmente são divididas por sexo.

A maioria das tábuas de vida começa com uma sequência de TEMs  ${}_nM_x$  e deriva o resto da tábua a partir daí. Entretanto, desde já é importante fazer duas observações a respeito:

1. A opção pela construção da tábua de vida a partir de  ${}_nM_x$  se baseia em considerações práticas, já que em muitos casos a informação estatística de base vem neste formato. Mas em realidade, as funções da tábua de vida são equivalentes, no sentido de que qualquer uma delas pode ser construída a partir de qualquer outra. Em teoria seria possível (embora pouco usual) construir uma tábua de vida inteira a partir da série de dados que descrevem o tempo médio que um indivíduo de idade exata  $x$  ( $x = 0, 1, 5, 10, 15, 20, 25$  etc. anos) ainda tem por viver (a função da tábua de vida conhecida como  $e_x$ )<sup>1</sup>. Existem outras situações, mais realistas, em que a tábua precisa ser construída a partir de outros elementos do que a sequência  ${}_nM_x$ . Por exemplo, em estudos de coorte, principalmente quando baseados em poucos indivíduos e informação completa sobre as suas idades de falecimento, faz mais sentido construir a tábua a partir da função  $\ell_x$ , que descreve o número de sobreviventes em cada idade  $x$ . Por outro lado, como foi indicado no Capítulo 8, o cálculo de  ${}_nM_x$  requer estatísticas fidedignas sobre o número de óbitos por sexo e idade e dados censitários para estimar os denominadores correspondentes. Atualmente Portugal, o Brasil, Macau, Cabo Verde<sup>2</sup> e eventualmente São Tomé & Príncipe são os únicos países lusófonos que dispõem destes dados de forma direta, por meio do Registro Civil ou do SIM. Nos outros países lusófonos, da mesma forma como noutros países em desenvolvimento, não há informação confiável sobre a sequência  ${}_nM_x$ , mas sim sobre outros indicadores que, com a ajuda de modelos, podem ser convertidos numa tábua de vida inteira. Estas situações serão exploradas em mais detalhe nos Capítulos 20 e 23. O que importa aqui é distinguir entre a lógica interna da tábua de vida – ou seja, a relação entre as diferentes funções – e a sequência da sua construção, que nem sempre é a mesma.
2. No caso em que a tábua é construída a partir de  ${}_nM_x$ , é importante entender que existe uma diferença entre  ${}_nM_x$  (a TEM na população observada) e  ${}_n m_x$  (a TEM que seria observada se a estrutura etária da população correspondesse exatamente à tábua de vida). Como será discutido

<sup>1</sup> A única situação em que isso às vezes ocorre é aquela em que a série  $e_x$  foi publicada, mas o resto da tábua de vida em que a série se baseia não. Por exemplo, o *Anuario Demográfico de Cuba 2001* (pág. 95) – talvez para poupar espaço – publicou uma série  $e_x$  completa por idade simples, mas omitiu o resto da tábua. Num caso como este, embora seja um cálculo pouco usual, é possível reconstruir a tábua toda a partir desta informação.

<sup>2</sup> Apesar de serem considerados relativamente completos, os dados de mortalidade foram corrigidos em 13,1% para os homens e 21,3% para as mulheres, para efeitos da construção de tábuas de vida na ocasião do Censo de 2010 (INE Cabo Verde, 2014: 33).

em mais detalhe na seção 9.3, numa população onde por muito tempo nasce o mesmo número de pessoas todos os anos e o nível de mortalidade não muda (uma chamada população *estacionária*), a estrutura etária corresponde exatamente à função de sobrevivência  $\ell_x$  da tábua de vida. As TEMs numa população deste tipo correspondem à sequência  ${}_n m_x$ . Mas noutras populações com a mesma tábua de vida a estrutura etária observada pode ser bastante diferente e isso afetará as TEMs em alguma medida, especialmente naqueles intervalos etários onde a intensidade da mortalidade varia muito. Por exemplo, na faixa de 1-4 anos a mortalidade em muitos países diminui acentuadamente com a idade. Uma criança de 1 ano de idade costuma correr um risco consideravelmente maior do que uma criança de 4 anos que já passou pela fase dos primeiros impactos ambientais e que já sobreviveu às primeiras agressões de agentes infecciosos. Sendo assim, a TEM para o intervalo de 1-4 anos ( ${}_4 M_1$ ) dependerá, até certo ponto, da distribuição das crianças neste intervalo. Se houver mais crianças de 1 ano do que crianças de 4 anos, a  ${}_4 M_1$  será mais alta; se a distribuição for contrária, a  ${}_4 M_1$  será mais baixa. Como a diferença entre  ${}_n m_x$  e  ${}_n M_x$  geralmente não é muito grande, a maioria das técnicas de construção de tábuas de vida mostradas aqui ignoram a distinção, mas também se mostrará uma técnica (o segundo método de Keyfitz) que a considera explicitamente.

Já foram mencionadas algumas das funções da tábua de vida. A seguir se mostra uma lista mais completa:

- ${}_n m_x$  A *Taxa Específica de Mortalidade* da tábua de vida (ou da população estacionária) do intervalo etário de  $x$  a  $x+n$  anos exatos (ou  $x$  a  $x+n-1$  anos completos).
- ${}_n M_x$  A *Taxa Específica de Mortalidade* correspondente na população observada, geralmente próxima mas não exatamente igual a  ${}_n m_x$ .
- ${}_n q_x$  A *probabilidade de morte* do intervalo etário de  $x$  a  $x+n-1$  anos completos, ou seja, a probabilidade de que uma pessoa de  $x$  anos exatos morra antes de alcançar a idade de  $x+n$  anos exatos. A terminologia “taxas do tipo q” já foi introduzida no Capítulo 7.
- ${}_n p_x$  O complemento de  ${}_n q_x$ , ou seja  ${}_n p_x = 1 - {}_n q_x$ , a probabilidade de sobrevivência do intervalo etário de  $x$  a  $x+n-1$  anos completos.
- $\ell_x$  O número de *sobreviventes* de uma coorte de  $\ell_0$  nascidos vivos depois de  $x$  anos<sup>3</sup>. No caso onde  $\ell_0 = 1$ , nota-se que  $\ell_x$  é equivalente com  ${}_x p_0$ , embora essa notação seja raramente usada.
- ${}_n d_x$  O *número de óbitos* esperado entre uma coorte de  $\ell_0$  nascidos vivos entre as idades de  $x$  e  $x+n-1$  anos completos.
- ${}_n L_x$  O número de *anos-pessoa* vivido pela coorte de  $\ell_0$  pessoas nascidas vivas e ainda vivas na idade de  $x$  anos no intervalo de  $x$  até  $x+n$  anos exatos. Se todo o mundo sobrevive dos  $x$  até os  $x+n$  anos, este número será simplesmente  $n \ell_x$ , mas como algumas pessoas morrem durante o intervalo, o número efetivo será menor:  $n \ell_{x+n} + {}_n a_x {}_n d_x$ , onde  ${}_n a_x$  é:

<sup>3</sup> Em realidade, se trata da letra “l”, de Leonardo, mas para evitar confusão com o número 1, optou-se aqui por usar a forma manuscrita  $\ell$ .

- ${}_n a_x$  Um *fator de separação* (terminologia introduzida em (7.6)) que expressa o número médio de anos vividos entre as idades  $x$  e  $x+n$  por pessoas vivas em  $x$ , mas que morrem antes de  $x+n$ . Normalmente este número é aproximadamente  $n/2$ , mas pode se afastar deste valor se há variações fortes da intensidade da mortalidade ao longo do intervalo.
- ${}_n S_x$  A *razão de sobrevivência* do intervalo etário  $(x, x+n)$  ou seja a proporção de pessoas com idades entre  $x$  e  $x+n$  anos exatos na população estacionária que sobreviverão até o próximo intervalo:  ${}_n L_{x+n} / {}_n L_x$ .
- $T_x$  O número de *anos-pessoa* vivido pela coorte de  $\ell_0$  pessoas nascidas vivas e ainda vivas na idade de  $x$  anos até a extinção completa da coorte, ou seja, até a morte do último sobrevivente.
- $e_x$  O *número médio de anos* vividos a partir da idade exata  $x$ . Em particular,  $e_0$  simboliza a *esperança de vida* (também chamada *expectativa de vida*) ao nascer.

Além destas funções padrão da tábua de vida, há outra que não faz parte da tábua como tal, mas que costuma ser útil do ponto de vista teórico, a saber:

- $\mu(x)$  A *força da mortalidade ou taxa instantânea de mortalidade* na idade  $x$ . A forma mais fácil de entender esta quantidade é como a Taxa de Mortalidade Específica na idade  $x$  para um intervalo de idade muito pequeno, tão pequeno que a distribuição da população ao longo do intervalo se torna irrelevante. Embora esta seja uma quantidade muito útil do ponto de vista conceitual, o seu manuseio exige um conhecimento básico de cálculo diferencial e integral, razão pela qual a sua discussão sistemática será adiada até o Capítulo 17.

É importante familiarizar-se com o significado de cada símbolo porque a partir deste capítulo se fará referência a eles com frequência. As próximas seções explicarão como calcular cada uma das funções da tábua de vida.

Embora a origem da tábua de vida esteja nos estudos da mortalidade humana, a mesma metodologia pode ser generalizada e estendida de várias maneiras. Desde os dias de Graunt e Halley a tábua de vida vem sendo utilizada para mensurar a esperança de vida não somente de populações humanas, mas de qualquer “coisa” que possua um tempo de vida útil e que possa ser acompanhada ao longo do tempo numa visão de coorte, ou cujo comportamento dentro de um determinado período possa ser observado para diferentes coortes. Para começar, a tábua de vida, principalmente a tábua de vida de coortes, é muito usada na ecologia da população, para descrever a sobrevivência de animais ou plantas. Mas, como já se mostrou na seção 3.10, também é possível utilizar a metodologia de tábuas de vida para estudar o ciclo de vida das empresas ou das moradias físicas (Gleeson, 2007) ou se pode utilizar tábuas de vida para representar a expectativa da duração média de casamentos ou o intervalo entre um nascimento e o próximo, ou ainda, é possível, a partir da construção de tábuas de vida, medir probabilidades de funcionários de um grande conglomerado empresarial, ou grande indústria, por aposentadoria, morte ou demissão (Tsai, Bernacki e Lucas, 1989; Hoverstad, Moncrief e Lucas, 1990). No Capítulo 12 se discutirá o método de Hajnal para determinar a idade média ao casar que, no fundo, nada mais é do que uma “esperança de vida solteira”. Nessa lógica, é conveniente abstrair o evento óbito (ou morte) como sendo a única forma

de saída de uma coorte. O conceito de *decremento* (em inglês também se usa o termo “attrition” ou *atrído*) parece ser o mais adequado quando se pretende generalizar as formas pelas quais alguém ou alguma coisa deixa de existir ou de ser observado durante um período.

Então, uma tábua de vida de decremento único é aquela em que só existe uma forma de saída da coorte em acompanhamento. No caso de uma tábua de vida de decremento único para populações humanas a única forma de saída é pela morte. Por outro lado, não é raro o interesse em tábuas de vida de múltiplos decrementos, mesmo quando o evento em observação é a morte. Por exemplo, uma morte por doenças cardiovasculares é um tipo de decremento, enquanto a morte por acidentes de trânsito é outro tipo de decremento. Na área de saúde pública e epidemiologia é grande o interesse por tábuas de vida por múltiplos decrementos com o objetivo de avaliar ganhos em termos de esperança de vida quando se reduz ou se elimina a incidência/prevalência de mortes por determinada causa básica. Assunto este que será discutido com maiores detalhes na seção 9.4.

Há ainda situações em que a metodologia de construção de uma tábua de vida pode ser generalizada para incorporar não somente uma ou várias formas de decremento, mas também uma ou várias formas de incremento. Nesse caso, a coorte em acompanhamento pode aumentar ou diminuir em volume não somente por saídas, mas também por entradas, como no caso de uma coorte de trabalhadores que ingressaram na indústria têxtil no ano 2000. Se permanecerem nesse mercado, os indivíduos permanecem na coorte, mas se saem por morte, invalidez, aposentadoria ou demissão, ou outra causa, deixam de pertencer a esta coorte. Contudo, empregados demitidos de uma empresa podem passar algum tempo fora do mercado de trabalho e depois voltar a ter um emprego no mesmo ramo. De forma similar, alguém que estava afastado por motivo de saúde pode retomar suas atividades, se reintegrando à coorte. Para analisar esse tipo de situação faz-se necessária uma tabela de incremento-decremento, pela qual se pode sair e voltar à coorte inicial.

Por suas diferentes aplicações, as tabelas de vida estão presentes em inúmeros estudos e análises, tendo aplicações em políticas públicas, como as da área de saúde, previdência, educação, desenvolvimento econômico, desenvolvimento social, entre outras, assim como aplicações na biologia, agronomia, economia, entre outras áreas.

## 9.2 TÁBUA DE VIDA DE COORTE (GERAÇÃO): DEFINIÇÕES BÁSICAS, FUNDAMENTOS E NOTAÇÕES

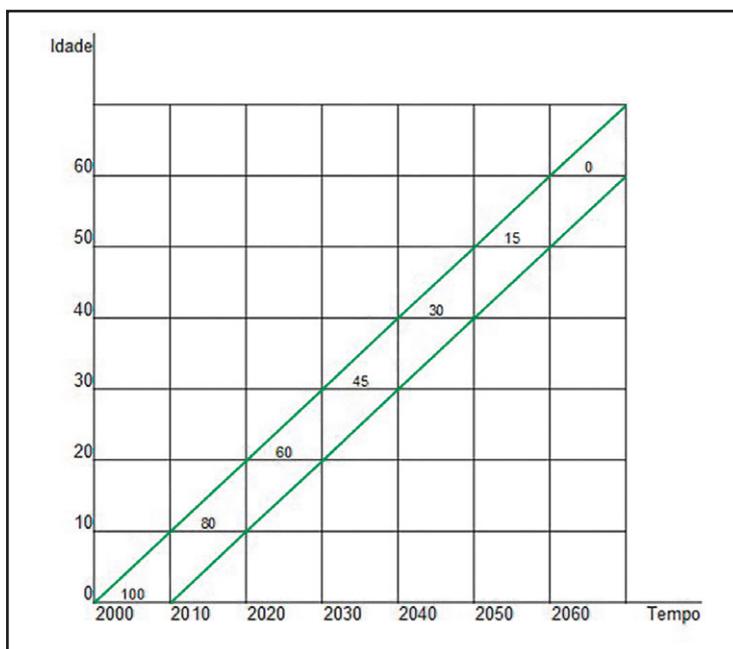
As funções na tábua de vida são expressas por idade simples (tábua completa) ou por grupos etários (tábua abreviada) e estão relacionadas de tal forma que, conhecendo-se uma única função, é possível determinar todas as demais. Talvez a forma mais simples de elucidar o processo de construção de uma tábua de vida seja como o acompanhamento de uma coorte ou uma geração, mediante o conceito de linhas vitais, tal como apresentado na seção 7.4. Suponha uma coorte de nascidos vivos num determinado ano  $t$ . Se fosse possível acompanhar essa coorte no tempo até a sua completa extinção existiriam todas as informações necessárias para exprimir toda a experiência de mortalidade dessa coorte no que se denomina a *tábua de vida de coorte*.

Como exemplo podem ser usados os dados apresentados no Gráfico 9.1. Neste exemplo, se observa uma coorte de 100 indivíduos nascidos entre 1º de janeiro do ano 2000 a 31 de dezembro do ano 2009 que sobrevivem a cada grupo decenal de idade a cada período de 10 anos. Mediante as definições e conceitos apresentados na seção 7.4, o diagrama do Gráfico 9.1 permite determinar o número de indivíduos desta coorte que sobrevive à idade exata  $x$  no início de cada grupo decenal

de idade (valores expostos nas linhas horizontais) ou que alcança com vida a data de 1º de janeiro no início de cada década (valores expostos nas linhas verticais). Neste exemplo com dados fictícios, o único decremento da coorte a ser observado é o óbito, ou seja, não há emigração ou imigração em qualquer grupo etário dessa coorte. Além disso, assume-se que todos os óbitos são devidamente registrados.

O diagrama expõe os dados de uma coorte de 100 nascidos vivos. Desta coorte, 20 indivíduos morreram antes de completar o décimo aniversário. Dos 80 indivíduos que sobreviveram à idade exata de 10 anos, apenas 60 sobreviveram à idade exata de 20 anos, e assim por diante. Assim, o diagrama representa o comportamento da coorte desde o nascimento até a sua completa extinção (nenhum indivíduo dessa coorte alcança com vida o sexagésimo aniversário).

Gráfico 9.1: Diagrama de Lexis de uma coorte hipotética nascida entre 1º de janeiro de 2000 e 31 de dezembro de 2010 acompanhada até 31 de dezembro de 2060

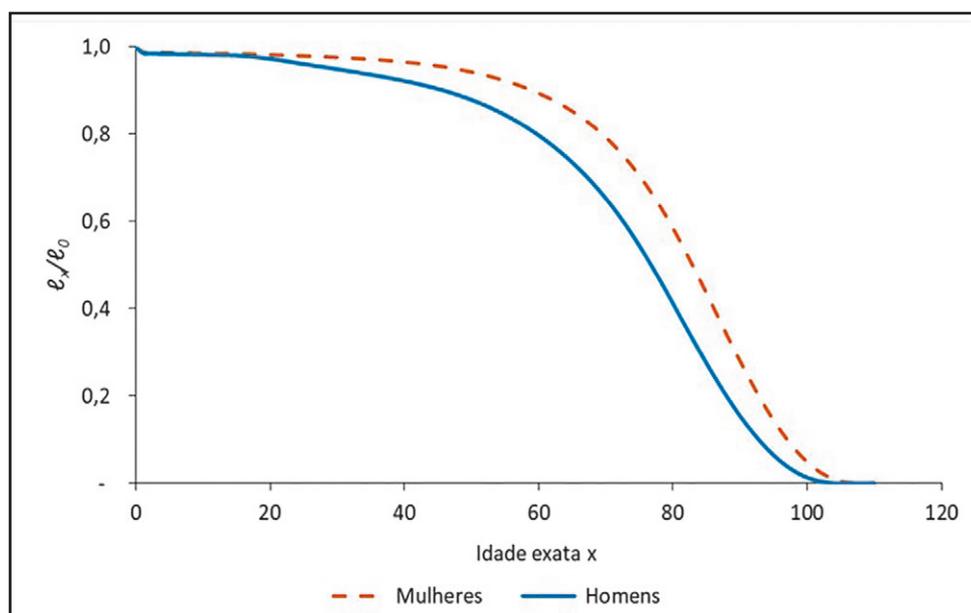


Os dados apresentados no diagrama do Gráfico 9.1 são suficientes para definir as principais funções de uma tábua de vida de coorte, a começar por  $\ell_x$ , o número de sobreviventes à idade exata  $x$ . No diagrama,  $\ell_0 = 100$ , é o tamanho inicial da coorte no decênio 2000-2010, ou seja, o número de indivíduos com idade 0 na data inicial;  $\ell_{10} = 80$  é o número de sobreviventes à idade exata de 10 anos e  $\ell_{20} = 60$  é o número de sobreviventes à idade exata de 20 anos.

Na tábua de vida, o valor  $\ell_0$  é denominado *raiz da tábua*. A raiz  $\ell_0$  pode ser qualquer valor, pois o que interessa é a relação entre os  $\ell_x$ , e não a escala de seu valor a cada idade. Ressalta-se que, como se trata do acompanhamento de uma coorte até sua extinção, o número de sobreviventes a cada idade exata  $x$ ,  $\ell_x$  é uma função positiva decrescente.

O Gráfico 9.2 apresenta a função  $\ell_x/\ell_0$  (equivalente com  ${}_x p_0$ , a probabilidade de sobreviver desde o nascimento até a idade exata  $x$ ) para homens e para mulheres elaborada pelo IBGE por

idade simples (grupos etários de 1 ano) de 0 a 110 anos considerando as funções de mortalidade observadas no Brasil em 2010<sup>4</sup>. Esse gráfico apresenta uma típica função  $\ell_x$ , em que se observa uma queda no valor de  $\ell_x$  nas primeiras idades devido à mortalidade infantil. Nas idades jovens e adultas a intensidade da queda de  $\ell_x$  é menor, pois o número de mortes nessas idades também é menor. Por fim, a inclinação da curva  $\ell_x$  volta a aumentar nas idades mais avançadas, pois aumenta a probabilidade de morte nessas idades. Repare ainda que a curva  $\ell_x$  dos homens tem níveis mais baixos que a das mulheres em função da sobremortalidade masculina.

Gráfico 9.2: Função  $\ell_x/\ell_0$  para homens e mulheres, Brasil, 2016

Fonte: IBGE (2016).

Se  $\ell_x$  pode diminuir com a idade, a função  ${}_n d_x$  define o número de mortes entre as idades exatas  $x$  e  $x+n$ . Portanto,

$${}_n d_x = \ell_x - \ell_{x+n} \quad (9.1)$$

Consequentemente  $\ell_{x+n} = \ell_x - {}_n d_x$ .

Voltando à coorte representada no Gráfico 9.1, o número de óbitos ocorridos entre 0 e 10 anos de idade,  ${}_{10}d_0$ , é 20, pois  $\ell_0 - \ell_{0+10} = 100 - 80 = 20$ . De forma similar, o número de óbitos ocorridos entre 10 e 20 anos de idade,  ${}_{10}d_{10}$ , é 20, pois  $\ell_{10} - \ell_{10+10} = 80 - 60 = 20$ .

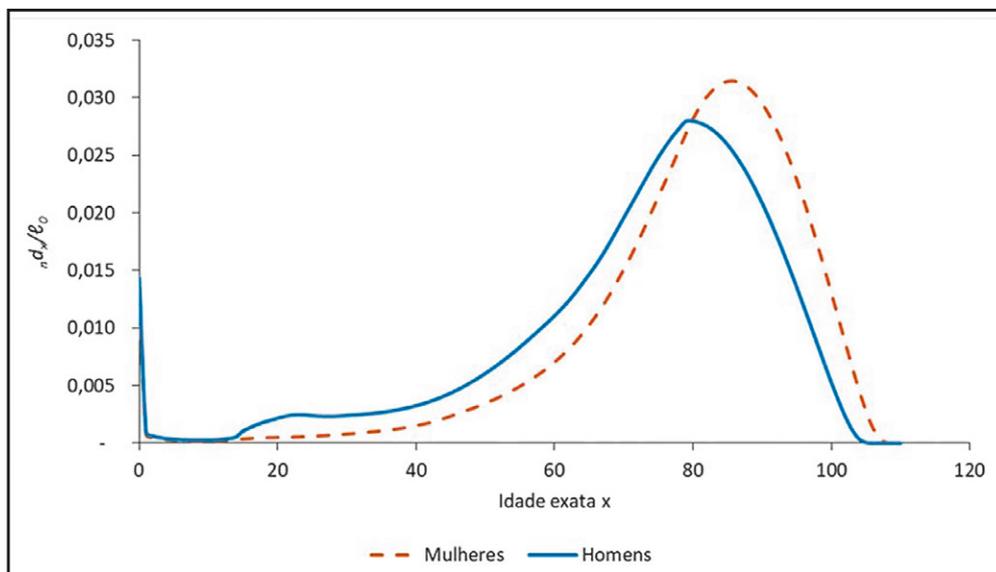
<sup>4</sup> Em realidade, a tábua de vida que serve de ilustração neste e outros gráficos mais adiante é uma tábua de período. Ela foi escolhida porque é difícil encontrar bons exemplos de tábuas de vida de coorte e porque a distinção não importa muito se o objetivo for apenas ilustrar o padrão etário das funções.

Uma dedução importante mediante a relação (9.1), tal como apontado por Ortega (1987), é que a soma dos óbitos a partir da idade  $x$  até a completa extinção da coorte na idade  $\omega$  é exatamente o valor de  $\ell_x$ , tal como segue:

$$\ell_x = \sum_{a=x}^{\omega-n} n d_a \quad (9.2)$$

O Gráfico 9.3 mostra a função  ${}_n d_x / \ell_0$  para a tábua de vida para o Brasil em 2010 quando  $n=1$ , ou seja, a proporção de mortes a cada idade em relação ao número de indivíduos iniciais. Como  $\ell_0$  é um valor arbitrário qualquer, o gráfico evidencia o formato típico de uma função  ${}_n d_x$ . Observe que esta função tem duas modas, a primeira nas primeiras idades, em função da maior fragilidade das crianças, e a segunda nas idades mais avançadas, como destacado por Ortega (1987). Quanto menor a mortalidade, mais avançada é a idade da segunda moda e menor a dispersão das mortes em torno desta idade. Como mostra o Gráfico 9.3, a mortalidade feminina ocorre com maior frequência em idades mais avançadas que a mortalidade masculina. Estudos mais recentes têm explorado o perfil etário da função  ${}_n d_x$  à direita da segunda moda, dado sua semelhança com a distribuição normal (Gaussiana) de probabilidade, para análise da idade modal, mediana ou média à morte, e também da variabilidade dos óbitos ao redor desta idade. Este tipo de análise tem sido um mecanismo útil para verificar as hipóteses relacionadas ao potencial de longevidade de uma população.

Gráfico 9.3: Função  ${}_n d_x / \ell_0$  para homens e mulheres, Brasil, 2016



Fonte: IBGE (2016).

Nos países desenvolvidos, uma das consequências do declínio histórico da mortalidade humana foi uma redução significativa na variabilidade da idade à morte (Wilmoth e Horiuchi, 1999). Define-se “variabilidade da idade à morte” como a dispersão da frequência de óbitos por idade ao redor de um valor médio, mediano ou modal da idade à morte dos indivíduos de uma

determinada população (Wilmoth e Horiuchi, 1999; Kannisto, 2001). Entender a variabilidade da idade à morte é importante uma vez que, em termos estatísticos, a distribuição dos óbitos ao longo das idades não é caracterizada somente pela média, mas também por medidas de dispersão. Juntas, essas medidas descrevem o padrão etário da mortalidade.

Historicamente, a redução na variabilidade da idade à morte nos países desenvolvidos ocorreu com base num declínio significativo da mortalidade nas idades mais jovens, principalmente da mortalidade infantil, devido à redução dos óbitos por causas exógenas. Em seguida, houve um processo de deslocamento das mortes para as idades adultas e mais avançadas, com uma redução na dispersão da distribuição dos óbitos por idade (processo denominado “compressão-retangularização”), devido, basicamente, a mudanças estruturais e progressos no combate às doenças não transmissíveis (Nusselder e Mackenbach, 1996; Wilmoth e Horiuchi, 1999; Edwards e Tuljapurkar, 2005; Oliveira, 2010). Quando proposto inicialmente por Fries (1980), esse o processo de compressão-retangularização seria um argumento a favor da existência de um limite biológico para a longevidade humana (Fries, 1980; Wilmoth, 1997). Com base numa inspeção visual da função de sobrevivência, existiria um argumento de que uma longevidade humana fixada em, por exemplo, 120 anos, faria com que a esperança de vida ao nascer não ultrapassasse os 85 anos e, conseqüentemente, a função de sobrevivência teria um formato cada vez mais retangular (Fries, 1980; Wilmoth, 1997). Atualmente, observações empíricas deste processo estão sendo utilizadas, analisadas também em países em desenvolvimento (Gonzaga, 2008; Gonzaga, Queiroz e Machado, 2008, 2009).

Da relação entre o número de mortes das idades  $x$  a  $x+n$ ,  ${}_n d_x$ , e o número de indivíduos expostos, inicialmente, ao risco de morrer nesse grupo etário,  $\ell_x$ , estima-se a probabilidade de morte entre as idades  $x$  e  $x+n$ ,  ${}_n q_x$ . Trata-se de uma probabilidade condicional de morte dado que um indivíduo sobreviveu à idade exata  $x$  que é determinada por:

$${}_n q_x = \frac{{}_n d_x}{\ell_x} = \frac{\ell_x - \ell_{x+n}}{\ell_x} \quad (9.3)$$

Conseqüentemente,  ${}_n d_x = \ell_x \cdot {}_n q_x$  e  $\ell_{x+n} = \ell_x \cdot (1 - {}_n q_x)$ .

Como toda probabilidade,  ${}_n q_x$  trata da relação entre número de eventos ocorridos (ou casos favoráveis) e número de pessoas expostas à ocorrência deste evento (ou casos possíveis), apresentando, portanto, valores entre 0 e 1. Uma definição alternativa para esta probabilidade é a de taxa inicial, tal como apresentada na seção 7.5.

Voltando à coorte de nascidos vivos entre 1º de janeiro de 2000 e 31 de dezembro de 2010 (Gráfico 9.1), constam os seguintes valores para  ${}_n q_x$  nos dois primeiros grupos etários decenais desta coorte:

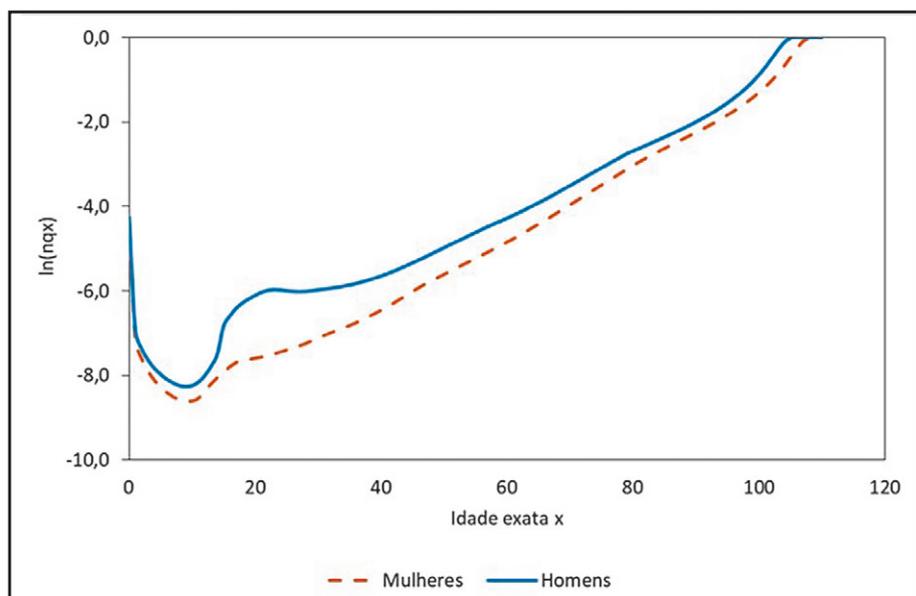
$$\begin{aligned} {}_{10}q_0 &= \frac{{}_{10}d_0}{\ell_0} = \frac{20}{100} = 0,200 \\ {}_{10}q_{10} &= \frac{{}_{10}d_{10}}{\ell_{10}} = \frac{20}{80} = 0,250 \end{aligned} \quad (9.4.a-b)$$

A função  ${}_nq_x$  para o Brasil é apresentada no Gráfico 9.4, numa escala logarítmica, para homens e mulheres, baseada no suposto de constância das taxas conforme o nível de 2010. Essa função, que tem formato conhecido como de “J”, evidencia a maior probabilidade de morte nas idades iniciais, caindo nas idades mais baixas e voltando a crescer com o avanço da idade, chegando a 1 no último grupo etário possível para a tabela construída. O formato de “J”, contudo, desaparece com a diminuição da mortalidade como um todo e, principalmente, com a diminuição da mortalidade infantil e na infância. A Islândia, que em 2016 alcançou uma mortalidade infantil de apenas 0,7 por mil nascimentos, a taxa mais baixa do mundo, foi o primeiro país a eliminar a sobremortalidade no primeiro ano de vida<sup>5</sup>.

Aparentemente o Gráfico 9.4 é idêntico ao Gráfico 8.1. Efetivamente, usando  $n=1$  a diferença entre  ${}_nM_x$  e  ${}_nq_x$  é tão pequena que quase não pode ser percebida num gráfico deste tipo. Em realidade,  ${}_nq_x$  é um pouco menor do que  ${}_nM_x$ , mas a diferença é mínima. Se fosse usado  $n=5$ , em vez de  $n=1$ ,  ${}_nq_x$  seria quase 5 vezes maior do que  ${}_nM_x$ .

Comparando as funções para homens e mulheres apresentadas no Gráfico 9.4 percebe-se que a função  ${}_nq_x$  evidencia a sobre mortalidade masculina nas idades entre 15 e 30 anos devido a causas externas, pois a função masculina apresenta valores mais elevados que a feminina nessas idades. Por ser de mais fácil interpretação visual, esta função é mais utilizada que as demais abordadas até agora. Epidemiologistas têm explorado uma variedade de modelos matemáticos para descrever o perfil etário desta função, especialmente em idades mais avançadas. Se destacam os modelos de Gompertz, Makeham e Heligman-Pollard que serão apresentados no Capítulo 20.

Gráfico 9.4: Função  ${}_nq_x$  para homens e mulheres, Brasil, 2016



Fonte: IBGE (2016).

<sup>5</sup> Entretanto, como a Islândia tem apenas 4.000-4.500 nascimentos por ano, a taxa varia consideravelmente de um ano para outro.

Enquanto  ${}_nq_x$  representa a probabilidade de morte entre duas idades exatas, dado que sobreviveu à idade inicial,  ${}_np_x$  representa a sobrevivência. Como numa tábua de vida só há duas possibilidades, estar vivo ou morto,  ${}_np_x$  é a probabilidade complementar de  ${}_nq_x$ :  ${}_np_x = 1 - {}_nq_x$ . Sendo  ${}_nq_x$  e  ${}_np_x$  probabilidades complementares, nota-se que no último grupo etário, enquanto a probabilidade de morrer é 1, já que se assume que todos morrem até o último grupo etário, a probabilidade de sobrevivência é 0.

Exemplificando com os dados do Gráfico 9.1:

$${}_{10}p_0 = 1 - {}_{10}q_0 = 0,800 \quad (9.5.a)$$

$${}_{10}p_{10} = 1 - {}_{10}q_{10} = 0,750 \quad (9.5.b)$$

Para poder calcular  ${}_nL_x$  e ultimamente a esperança de vida  $e_x$ , partindo de dados agrupados e não individuais, precisa-se da função  ${}_na_x$ , o tempo médio vivido no intervalo etário pelos  ${}_nd_x$  indivíduos que morreram entre as idades  $x$  e  $x+n$ . O valor exato dessa função é importante não só para determinar com mais exatidão as demais funções da tábua de vida, mas esta função isoladamente também pode ser utilizada, por exemplo, por atuários. Um exemplo disso é que planos previdenciários pagam benefícios mensais a indivíduos que sobrevivem até aquela data. Então, além de saber a probabilidade de os indivíduos receberem ou deixarem de receber benefícios a cada ano, é importante saber por quantos meses os que morrem em determinado ano irão receber benefícios naquele ano para determinar, com mais exatidão, o montante de recursos que devem estar disponíveis para pagar tais benefícios.

Alguns dos  ${}_nd_x$  indivíduos que morreram podem ter morrido em momentos próximos ao início do grupo etário e outros indivíduos podem ter morrido em momentos mais próximos ao final do grupo etário. Se as idades exatas à morte forem conhecidas, pode-se calcular exatamente qual foi o tempo vivido por cada indivíduo no grupo etário e, no conjunto de todos os indivíduos, qual foi o tempo médio vivido por todos que morreram naquela faixa etária.

Se a idade exata à morte não é conhecida pode-se recorrer a algumas aproximações para estimar  ${}_na_x$ . Entre as interpolações mais conhecidas estão a interpolação linear, pela qual se assume uma distribuição uniforme das mortes; a interpolação exponencial, que assume que  ${}_np_x$  é exponencial dentro do intervalo etário; e a interpolação harmônica, ou o suposto de Balducci, que assume que a função  ${}_np_x$  é harmônica no intervalo etário entre  $x$  e  $x+n$ , ou seja:

$${}_sp_x = \frac{1}{\left(1 - \frac{s}{n}\right) + \frac{s}{n} {}_np_x} \quad (9.6)$$

Usando uma interpolação exponencial, se obtém o seguinte resultado para  ${}_na_x$ :

$${}_na_x = -n \left( \frac{{}_np_x}{1 - {}_np_x} + \frac{1}{\ln({}_np_x)} \right) \quad (9.7)$$

enquanto a interpolação com (9.6) dá

$${}_n a_x = -n \left( \frac{{}_n p_x}{1 - {}_n p_x} + \frac{{}_n p_x \ln({}_n p_x)}{(1 - {}_n p_x)^2} \right) \quad (9.8)$$

Ambos os resultados não são difíceis de provar, mas como requerem o uso de integração, as derivações não são mostradas aqui.

Pela interpolação linear assume-se que as mortes são uniformemente distribuídas entre as idades  $x$  e  $x+n$ , ocorrendo, em média, na metade do intervalo etário, o que é uma suposição razoável para idades intermediárias da tábua de vida, sobretudo quando o intervalo etário é pequeno<sup>6</sup>. Assim,  ${}_n a_x = n/2$ . Contudo, esta suposição não é razoável para os grupos etários iniciais e finais da tábua de vida onde a mortalidade varia de uma forma mais acentuada. Entre as crianças, se a probabilidade de morte é maior para crianças mais novas, com grande diferença no valor dessa função a cada idade, pode-se supor que dentro de um mesmo grupo etário a variação também seja grande, não cabendo a pressuposição de uniformidade da distribuição das mortes. Nos últimos grupos etários valeria o mesmo raciocínio.

Existem várias alternativas adicionais para estimar  ${}_n a_x$ , algumas das quais serão mostradas na seção 9.4.

A próxima função da tábua de vida de coorte a ser definida é o número total de anos-pessoa vividos pela coorte entre as idades  $x$  e  $x+n$ ,  ${}_n L_x$ . Mais precisamente,  ${}_n L_x$  refere-se ao número de pessoas-ano que  $\ell_x$  indivíduos viverão, em conjunto, até a idade exata  $x+n$ , considerando o tempo vivido tanto pelos que morreram no intervalo etário ( ${}_n d_x \cdot a_x$ ) quanto pelos  $\ell_{x+n}$  que sobreviveram os  $n$  anos no intervalo ( $n \cdot \ell_{x+n}$ ).

$${}_n L_x = {}_n d_x \cdot a_x + n \ell_{x+n} = {}_n a_x \ell_x + (1 - {}_n a_x) \ell_{x+n} \quad (9.9)$$

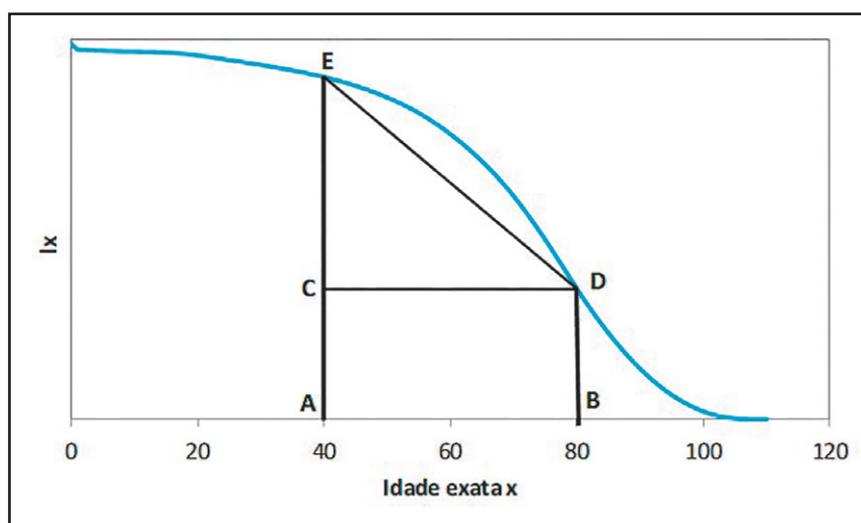
Como foi exposto acima, a forma mais simples de calcular  ${}_n L_x$  é assumindo a aproximação linear para  $\ell_x$ , com  ${}_n a_x = n/2$ . Alguns dados do Gráfico 9.1 podem ser usados para exemplificar o que isso significa. Da coorte inicial  $\ell_0 = 100$ , apenas  $\ell_{10} = 80$  pessoas sobreviveram à idade de 10 anos exatos. Destes,  $\ell_{20} = 60$  chegaram vivos à idade exata de 20 anos. Ou seja, há um total de  $\ell_{10} - \ell_{20} = 80 - 60 = 20$  óbitos entre as idades exatas de 10 e 20 anos. Portanto, o número de anos-pessoa vividos pela coorte entre as idades exatas de 10 e 20 anos é determinado pelas 60 pessoas que viveram 10 anos no intervalo entre as idades de 10 e 20 anos, somando 600 anos-pessoa, pois sobreviveram a todo o intervalo, mais o tempo vivido pelos 20 indivíduos que morreram em algum tempo durante o intervalo etário, cujo tempo médio vivido por cada um foi  ${}_{10} a_{10}$ . Assumindo que cada pessoa que morreu no intervalo de 10 anos viveu, em média, por um tempo igual à metade

<sup>6</sup> Este tipo de comportamento da função  $\ell_x$  é conhecido entre os atuários como a Lei de De Moivre. Abraham de Moivre, no seu livro, *Annuities upon Lives* (1725), estipulou que a função de sobrevivência  $\ell_x$  pode ser representada como  $\ell_x = \omega - x$  ( $0 \leq x \leq \omega$ ). Evidentemente, isso não é correto, mas a expressão tem certa importância histórica por ter sido a primeira representação formal proposta para a mortalidade humana e, por sua simplicidade, é conveniente para ilustrar certos princípios gerais da tábua de vida.

deste intervalo, ou seja, viveu por 5 anos, o total de anos-pessoa vividos por aqueles que morreram entre 10 e 20 anos é de  $20 \cdot 5 = 100$  anos. Portanto, a aproximação linear para  ${}_{10}L_{10}$  é de  $600 + 100 = 700$  anos. Pelo exposto no exemplo acima, se observa que cada indivíduo da coorte na idade exata  $x$ , que sobrevive para a idade exata  $x+n$ , contribui  $n$  anos para o cálculo de  ${}_nL_x$ , ao passo que cada indivíduo que morre entre  $x$  e  $x+n$  contribui apenas uma fração de  $n$  para  ${}_nL_x$ .

O cálculo mediante a aproximação linear também pode ser visualizado graficamente utilizando o perfil etário da função  $\ell_x$ , tal como mostrada no Gráfico 9.2. Geometricamente, o número de anos-pessoa vividos entre  $x$  e  $x+n$  é dado pelo valor da área delimitada pela função  $\ell_x$  e pelos eixos verticais em  $x$  e  $x+n$ , conforme apresentado no Gráfico 9.5. Nesse gráfico, a função  $\ell_x$  é uma função contínua em  $x$ . Ainda que fosse possível determinar  $\ell_x$  numa escala de tempo contínua, o que significa conhecer a função matemática de  $\ell_x$ , a determinação exata dessa área necessita de conhecimentos em cálculo diferencial e integral para funções contínuas. Porém, é possível obter cálculos aproximados bastante satisfatórios para valores inteiros de  $x$ , bastando determinar a área da figura geométrica regular delimitada por cada valor inteiro de  $x$ . No caso apresentado no Gráfico 9.5, é preciso determinar as áreas do retângulo ABCD e do triângulo ECD ou, alternativamente, do trapézio ABDE.

Gráfico 9.5: Anos-pessoa vividos por intervalo etário ( $x, x+n$ ) numa coorte hipotética



Fonte: IBGE (2016).

Na função  $\ell_x$  do Gráfico 9.5 a área delimitada pelos eixos verticais  $\ell_x$  e  $\ell_{x+n}$ , e que forma o retângulo ABCD, representa o número de anos-pessoa vividos por aqueles que alcançaram a idade  $x$  e sobreviveram à idade  $x+n$ . O resultado é a área de um retângulo cuja altura é o valor de  $\ell_{x+n}$  e a base é o tamanho do intervalo etário (aqui  $n = 80 - 40 = 40$  anos),  $\text{área} = n \cdot \ell_{x+n}$ . Já o total de anos-pessoa vividos pelos que morreram entre  $x$  e  $x+n$  é a área abaixo da função  $\ell_x$  delimitada pelos segmentos CE e CD. Novamente, o cálculo exato dessa área necessitaria do conhecimento da forma matemática da função  $\ell_x$  e de técnicas de integração para o cálculo de áreas não regulares.

Entretanto, fazendo uso da suposição de que o tempo ao óbito se distribui uniformemente dentro do intervalo  $(x, x+n)$ ,  $\ell_x$  decresce por uma variação constante formando uma reta que liga

os pontos E e D. Sob esta hipótese, o problema fica reduzido ao cálculo da área delimitada pelos pontos CDE (área de um triângulo retângulo =  $base \cdot altura / 2$ ). Neste triângulo CDE, o segmento CE representa a parcela de  $\ell_x$  que não sobreviverá à idade exata de  $x+n$  anos, ou seja, são os óbitos que ocorrerão entre  $x$  e  $x+n$  ( ${}_n d_x$ ). A área do triângulo CDE =  ${}_n d_x \cdot n / 2$  é uma estimativa para o total de anos-pessoa vividos pelos que morreram entre  $x$  e  $x+n$ , sob a hipótese de que os óbitos ocorreram uniformemente entre  $x$  e  $x+n$ , e sendo  $n/2$  o tempo médio vivido entre  $x$  e  $x+n$  por cada pessoa que morreu neste mesmo intervalo, segundo a aproximação linear. Logo, a soma das áreas dos triângulos ABCD e CDE é uma estimativa para  ${}_n L_x$  sob a hipótese de uniformidade (homogeneidade) do tempo ao óbito entre  $x$  e  $x+n$  e pode ser formalizada pela relação:

$${}_n L_x = n \ell_{x+n} + n (\ell_x - \ell_{x+n}) / 2 = n (\ell_x + \ell_{x+n}) / 2 \quad (9.10)$$

Aplicando esta lógica aos dados do Gráfico 9.1 confirma o resultado obtido acima, ou seja,

$$(80 + 60) = 5 \cdot 140 = 700 \quad (9.11)$$

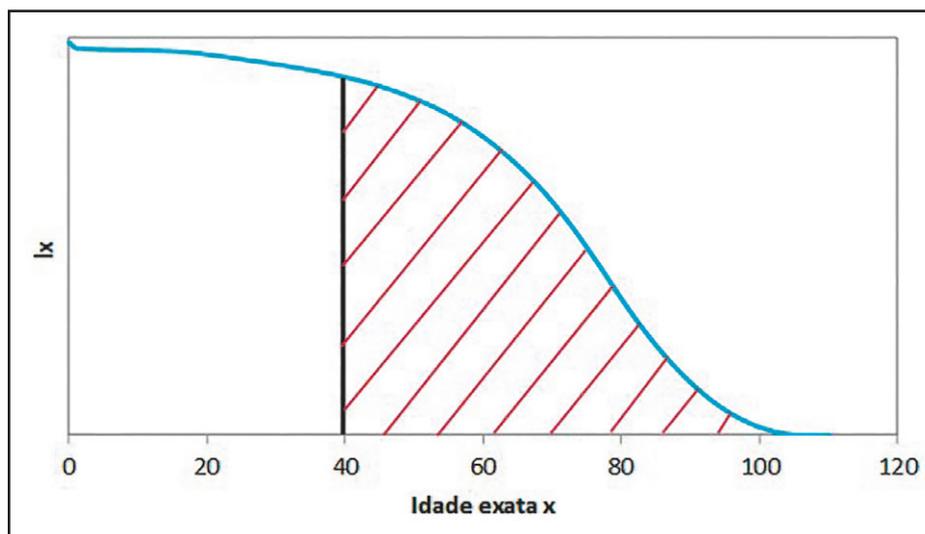
O intervalo etário (40,80) no Gráfico 9.5 evidentemente é muito grande e mesmo o intervalo de 10 anos em (9.11) é maior do que o normal. Isso acentua as distorções decorrentes do uso da aproximação linear, como pode ser visto claramente no gráfico. Usando intervalos convencionais de 5 anos ou, ainda mais, 1 ano, o problema é menor. Mesmo assim, se houver necessidade de uma aproximação mais próxima, pode-se recorrer a (9.7) ou (9.8), ou ainda a algumas das que serão expostas na seção 9.4.

A penúltima função da tábua de vida,  $T_x$ , é definida como o número total de anos vividos pela coorte a partir da idade  $x$ . Ela pode ser obtida pela soma de  ${}_n L_x$  a partir da idade  $x$  até a completa extinção da coorte:

$$T_x = \sum_{i=x}^{\infty} {}_n L_i \quad (9.12)$$

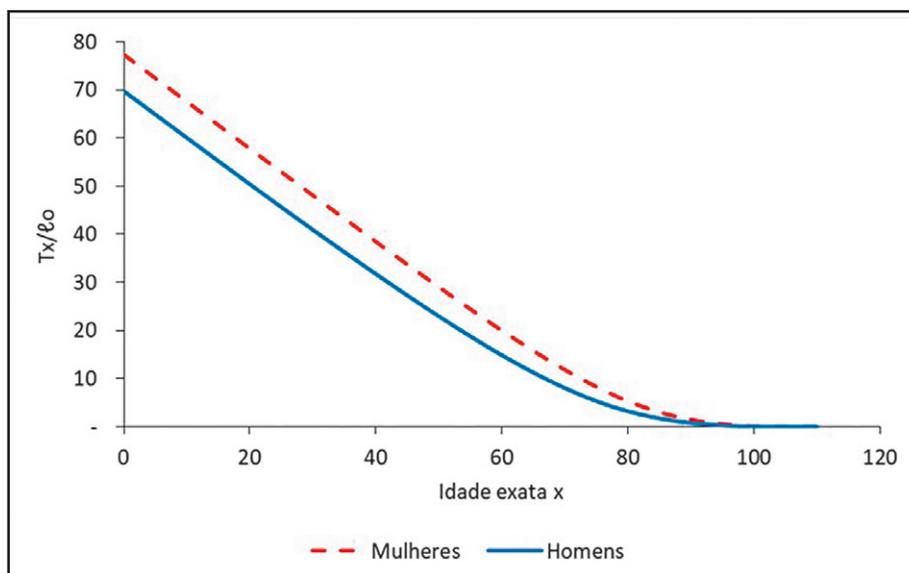
Pelos dados do Gráfico 9.1 obtém-se os valores de  ${}_n L_x$  apresentados na Tabela 9.1. Para a idade exata 10 anos,  $T_{10} = 700 + 525 + 375 + 225 + 75 = 1900$ .

Por (9.12) observa-se que se  $i = 0$  a função  $T_x$  representa o tempo total de anos vividos pela coorte desde o seu nascimento até a sua completa extinção. Assim como foi feito para a função  ${}_n L_x$ , se pode utilizar o perfil etário da função  $\ell_x$  para definir a área sob a função  $\ell_x$  que determina a função  $T_x$ , tal como mostrado no Gráfico 9.6. Assim como observado no Gráfico 9.5, se a área delimitada por  $\ell_x$  entre as idades  $x$  e  $x+n$  determinam o tempo total a ser vivido pela coorte entre as idades  $x$  e  $x+n$ , concluímos que toda a área sob a função  $\ell_x$  fornece o tempo total a ser vivido pela coorte desde o seu nascimento até a sua completa extinção, o que nos leva à definição para a função  $T_x$  quando  $x = 0$ . Se o interesse for definir  $T_x$  somente a partir da idade  $i = x$ , dá para notar que a área à direita da idade exata  $x$ , delimitada pela função  $\ell_x$ , tal como mostrado no Gráfico 9.6, fornece o tempo total a ser vivido a partir da idade  $x$ , por todas as pessoas da coorte que chegaram vivos à idade  $x$ .

Gráfico 9.6: Tempo vivido pela coorte a partir da idade  $x$  a partir dos 40 anos, homens, Brasil, 2016

Fonte: IBGE (2016).

Já o Gráfico 9.7 apresenta uma típica função  $T_x$  por idade e sexo, com base nos dados da tábua de vida do IBGE para o Brasil em 2010. Como essa função representa a soma dos tempos vividos a partir da idade  $x$ , essa função é não crescente para qualquer idade  $x$ . Mais uma vez, como mostra o gráfico, a função estudada apresenta maiores valores para mulheres que para homens em função da maior sobrevida feminina.

Gráfico 9.7: Tempo vivido pela coorte a partir da idade  $x$  ( $T_x$ ) por idade e sexo, Brasil, 2016

Fonte: IBGE (2016).

A última função da tábua de vida é a esperança de vida à idade exata  $x$ , ou seja, o tempo médio a ser vivido por determinada coorte a partir daquela idade. Para construir uma medida, que reflete

o tempo médio a ser vivido pela coorte a partir da idade  $x$ , é preciso retomar o conceito da média como medida de tendência central de uma distribuição de valores, ou seja, aquela que reflete um ponto de equilíbrio dessa distribuição (ou do conjunto de valores).

Na presente situação, pode-se pensar no conjunto de linhas vitais (ver definição na seção 7.4) de uma coorte. Suponha, então, a coorte de 100 nascidos vivos ( $\ell_0$ ) do Gráfico 9.1. Supõe-se que cada indivíduo da coorte viverá por, exatamente, 100 anos. Ou seja, a função  $\ell_x$  será definida como uma reta paralela ao eixo da idade e seu valor será 100 para qualquer idade  $x$  variando de 0 a 100. Existe, então, um conjunto de 100 linhas vitais medindo, exatamente, 100 anos cada. Noutras palavras, há um total de 10.000 anos-pessoa a serem vividos por esta coorte de 100 nascidos vivos. Nesta situação hipotética  $T_0 = 10.000$ . Este valor para  $T$  em  $x = 0$  é simplesmente a área delimitada pela função constante  $\ell_x = 100$ , para  $x$  variando de 0 a 100 anos. Assim, o tempo médio a ser vivido por essa coorte desde o seu nascimento ou, ou seja, a esperança de vida ao nascer da coorte, será exatamente de 100 anos, pois não há variabilidade no tempo vivido por cada indivíduo dessa coorte. A conta feita aqui é a de uma média aritmética simples:

$$\frac{\text{Tempo total a ser vivido pela coorte desde o nascimento}}{\text{Total de nascidos vivos}} = \frac{10.000}{100} = 100 \text{ anos} \quad (9.13)$$

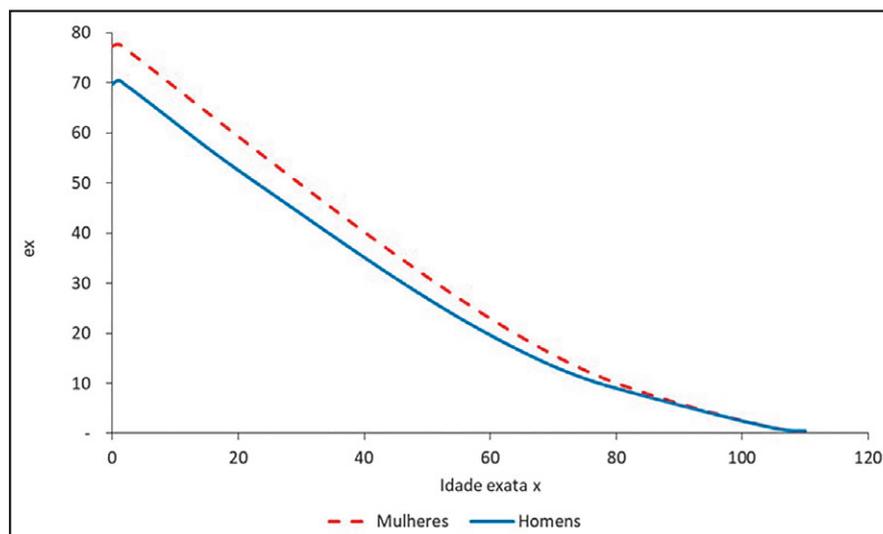
Conclui-se, formalmente, que a esperança de vida à idade  $x$  será dada por:

$$e_x = T_x / \ell_x \quad (9.14)$$

A relação (9.14) toma a razão entre o tempo total a ser vivido pela coorte a partir da idade  $x$  pelo número de indivíduos que chegaram vivos à idade  $x$ , o que leva ao conceito de média tal como visto no exemplo anterior. Em particular, se  $x = 0$ , tem-se a esperança de vida ao nascer.

O Gráfico 9.8 apresenta a esperança de vida por idade e sexo para o Brasil em 2010. Pelas estimativas do IBGE a esperança de vida feminina é maior que a masculina em todas as idades, sendo essa diferença maior para as primeiras idades que para as mais avançadas. Isso se dá porque a esperança de vida é uma medida síntese de toda a experiência de mortalidade a partir de cada idade. Como há grandes diferenças entre a mortalidade de homens e mulheres no decorrer das idades, essas diferenças se refletem na esperança de vida desde as idades mais jovens. Nota-se, ainda, que a esperança de vida cai com a idade para praticamente todas as idades, com exceção das primeiras. Nos primeiros anos de vida a probabilidade de morte é alta. A criança que morre deixa de contribuir com anos-pessoa para todas as demais idades. Portanto, a morte de um recém-nascido tem grande peso sobre o tempo de vida médio a ser vivido. Mas, uma vez superado este risco, é mais provável que os indivíduos permaneçam vivos e contribuam com linhas de vida mais extensas, o que leva ao aumento da esperança de vida. Por conta deste efeito, diz-se que variações na mortalidade infantil são as que têm maior efeito na esperança de vida de uma população.

Gráfico 9.8: Esperança de vida por sexo e idade, Brasil, 2016



Fonte: IBGE (2016).

Finalmente, uma quantidade diretamente associada a  ${}_nL_x$  é a razão de sobrevivência  ${}_nS_x$  que quantifica a proporção das pessoas que atualmente têm entre  $x$  e  $x+n$  anos de idade e que sobreviverão por mais  $n$  anos para compor a faixa etária de  $x+n$  a  $x+2n$  anos naquele momento. A fórmula para esta quantidade é simplesmente

$${}_nS_x = {}_nL_{x+n} / {}_nL_x \quad (9.15)$$

A utilidade principal de  ${}_nS_x$  surge na metodologia de projeções demográficas por componentes (ver Capítulo 21), onde é preciso saber justamente quantas pessoas com idades entre  $x$  e  $x+n$  anos continuarão fazendo parte da população em  $n$  anos.

A Tabela 9.1 apresenta um exemplo de uma tábua de vida de coorte, no caso a coorte de nascidos vivos em Portugal em 1940-44. Encontrar bons exemplos de tábuas de vida de coorte não é fácil porque requer estatísticas de mortalidade de boa qualidade durante um período de pelo menos 70 anos no passado e de preferência mais. Os principais exemplos de dados deste tipo são para países com uma longa história de boas estatísticas demográficas, tais como França, Países Baixos, Inglaterra, Itália, Nova Zelândia, Suíça e os países nórdicos. No caso de Portugal, as estatísticas de mortalidade anteriores a 1940 não foram consideradas suficientemente confiáveis pela Human Mortality Data Base que serviu de base para a Tabela 9.1, de modo que a coorte de 1940-44 foi a primeira que permitia a construção de uma tábua de vida como a mostrada aqui. Como resultado, a tábua é relativamente curta, terminando aos 70 anos, quando normalmente seria desejável continuar pelo menos até os 85. As esperanças de vida aos 70 anos (15,7 anos para homens e 19,9 para mulheres) não podem ser obtidas diretamente dos dados mas tiveram que ser estimadas a partir de outras tábuas de vida, com comportamentos semelhantes até os 70 anos.

Um outro problema dos dados portugueses é que Portugal sofreu volumes consideráveis de migração internacional, particularmente com o retorno de muitos portugueses das colônias em

meados dos anos 70. Por esta razão, a sequência  $N_x$ , que retrata a evolução do número de pessoas nascidas em 1940-44 ao longo do tempo, não corresponde ao número real de pessoas nas idades apropriadas presentes na população em cada ano. A sequência tal como aparece na Tabela 9.1 exclui os movimentos migratórios ocorridos entre 1940 e a atualidade, tanto a imigração como a emigração.

Mesmo com estas deficiências, a Tabela 9.1 dá uma boa ideia de como funciona a mecânica da construção de uma tábua de vida de coorte. Trata-se de uma planilha que pode ser montada facilmente em EXCEL. Se a primeira linha for 1 e a primeira coluna A e se a coluna B contiver os valores dados de  $N_x$  e a coluna D os multiplicadores  ${}_n a_x$ , as demais colunas podem ser obtidas da seguinte forma:

$$\ell_0 \text{ até } \ell_{70}: \quad C2 = 100000 * B1 / B\$1 \text{ e assim para baixo até a linha 17}$$

$${}_1 L_0: \quad E2 = D2 * C2 + (1 - D2) * C3$$

$${}_4 L_1: \quad E3 = D3 * C3 + (4 - D3) * C4$$

$${}_5 L_5 \text{ até } {}_5 L_{65}: \quad E4 = D4 * C4 + (5 - D4) * C5 \text{ e assim para baixo até a linha 16 (relação 9.9)}$$

$${}_{5+} L_{70}: \quad E17 = K17 * C17$$

$${}_1 d_0 \text{ até } {}_{5+} d_{70}: \quad F2 = C2 - C3 \text{ e assim para baixo (relação 9.1)}$$

$${}_1 p_0 \text{ até } {}_{5+} p_{70}: \quad G2 = C3 / C2 \text{ e assim para baixo}$$

$${}_1 q_0 \text{ até } {}_{5+} q_{70}: \quad H2 = 1 - G2 \text{ e assim para baixo (relação 9.3)}$$

$$T_{70}: \quad J17 = E17$$

$$T_0 \text{ até } T_{65}: \quad J16 = J17 + E16 \text{ e assim para cima (relação 9.12)}$$

$$e_0 \text{ até } e_{65}: \quad K2 = J2 / C2 \text{ e assim para baixo até a linha 16 (relação 9.14)}$$

(NB:  $e_{70}$  é um número dado)

$${}_1 m_0 \text{ até } {}_{5+} m_{70}: \quad I2 = F2 / E2 \text{ e assim para baixo}$$

Em teoria, esta última função ( ${}_n m_x = {}_n d_x / {}_n L_x$ ) não é estritamente necessária na construção de uma tábua de vida por coorte, razão pela qual ela não foi discutida antes, embora seja costumeiro incluí-la. Entretanto, como se verá na próxima seção, ela é fundamental na construção de tábuas de vida de período.

Tabela 9.1: Exemplo de tábua de vida de coorte: nascidos vivos em Portugal (1940-44)

	A	B	C	D	E	F	G	H	I	J	K
<b>1</b>	<b>Homens</b>	<b>P<sub>x</sub></b>	<b>l<sub>x</sub></b>	<b>n<sub>ax</sub></b>	<b>nL<sub>x</sub></b>	<b>nd<sub>x</sub></b>	<b>np<sub>x</sub></b>	<b>nq<sub>x</sub></b>	<b>nm<sub>x</sub></b>	<b>T<sub>x</sub></b>	<b>e<sub>x</sub></b>
2	0	497203	100000	0,3413	90836	13913	0,8609	0,1391	0,153166	6002049	60,02
3	1-4	428028	86087	1,8	328635	7143	0,9170	0,0830	0,021734	5911214	68,67
4	5-9	392515	78945	2,5	392520	881	0,9888	0,0112	0,002245	5582579	70,72
5	10-14	388133	78063	2,5	389350	387	0,9950	0,0050	0,000993	5190059	66,49
6	15-19	386211	77677	2,5	387187	479	0,9938	0,0062	0,001236	4800709	61,80
7	20-24	383832	77198	2,5	384508	593	0,9923	0,0077	0,001543	4413522	57,17
8	25-29	380882	76605	2,51	381307	690	0,9910	0,0090	0,001809	4029014	52,59
9	30-34	377452	75915	2,52	377509	833	0,9890	0,0110	0,002207	3647707	48,05
10	35-39	373310	75082	2,53	372967	989	0,9868	0,0132	0,002651	3270198	43,56
11	40-44	368394	74093	2,54	367183	1335	0,9820	0,0180	0,003635	2897231	39,10
12	45-49	361757	72758	2,55	359424	1783	0,9755	0,0245	0,004961	2530048	34,77
13	50-54	352892	70975	2,56	348839	2475	0,9651	0,0349	0,007094	2170625	30,58
14	55-59	340588	68501	2,57	334679	3220	0,9530	0,0470	0,009621	1821786	26,60
15	60-64	324578	65281	2,58	316450	4113	0,9370	0,0630	0,012998	1487107	22,78
16	65-69	304127	61168	2,59	293128	5274	0,9138	0,0862	0,017992	1170657	19,14
17	70+	277905	55894	2,6	877529	55894	0,0000	1,0000	0,063694	877529	15,70
<b>18</b>											
<b>19</b>	<b>Mulheres</b>	<b>P<sub>x</sub></b>	<b>l<sub>x</sub></b>	<b>n<sub>ax</sub></b>	<b>nL<sub>x</sub></b>	<b>nd<sub>x</sub></b>	<b>np<sub>x</sub></b>	<b>nq<sub>x</sub></b>	<b>nm<sub>x</sub></b>	<b>T<sub>x</sub></b>	<b>e<sub>x</sub></b>
20	0	462002	100000	0,3413	91779	12481	0,8752	0,1248	0,135985	6768430	67,68
21	1-4	404342	87519	1,8	334671	7003	0,9200	0,0800	0,020925	6676651	76,29
22	5-9	371988	80516	2,5	400584	799	0,9901	0,0099	0,001995	6341980	78,77
23	10-14	368295	79717	2,5	397788	319	0,9960	0,0040	0,000803	5941396	74,53
24	15-19	366820	79398	2,5	396276	285	0,9964	0,0036	0,00072	5543608	69,82
25	20-24	365502	79113	2,5	394787	310	0,9961	0,0039	0,000786	5147332	65,06
26	25-29	364068	78802	2,51	393118	359	0,9955	0,0045	0,000912	4752545	60,31
27	30-34	362412	78444	2,52	391208	408	0,9948	0,0052	0,001042	4359427	55,57
28	35-39	360528	78036	2,53	388923	509	0,9935	0,0065	0,001309	3968219	50,85
29	40-44	358176	77527	2,54	385930	693	0,9911	0,0089	0,001796	3579296	46,17
30	45-49	354974	76834	2,55	381887	931	0,9879	0,0121	0,002439	3193366	41,56
31	50-54	350671	75902	2,56	376407	1273	0,9832	0,0168	0,003381	2811479	37,04
32	55-59	344791	74630	2,57	369320	1576	0,9789	0,0211	0,004267	2435072	32,63
33	60-64	337511	73054	2,58	360288	2058	0,9718	0,0282	0,005713	2065753	28,28
34	65-69	328001	70996	2,59	348245	2794	0,9607	0,0393	0,008022	1705464	24,02
35	70+	315094	68202	2,6	1357219	68202	0,0000	1,0000	0,050251	1357219	19,90

Fonte: Calculado a partir da Human Mortality Data Base com valores estimados para  $e_{70}$ .

### 9.3 TÁBUA DE VIDA E POPULAÇÃO ESTACIONÁRIA

Como preparação para a discussão da tábua de vida de período na próxima seção vale a pena parar um momento para introduzir o conceito de população estacionária. Uma população é *estacionária* se as suas taxas de mortalidade e fecundidade por idade não variarem, se o número de óbitos for exatamente igual ao número de nascimentos e se não houver migração internacional. Em realidade, a população é um caso especial da população *estável*, um conceito mais geral que será discutido no Capítulo 22, na qual pode haver uma diferença entre a número de óbitos e nascimentos, desde que a diferença entre as taxas for constante.

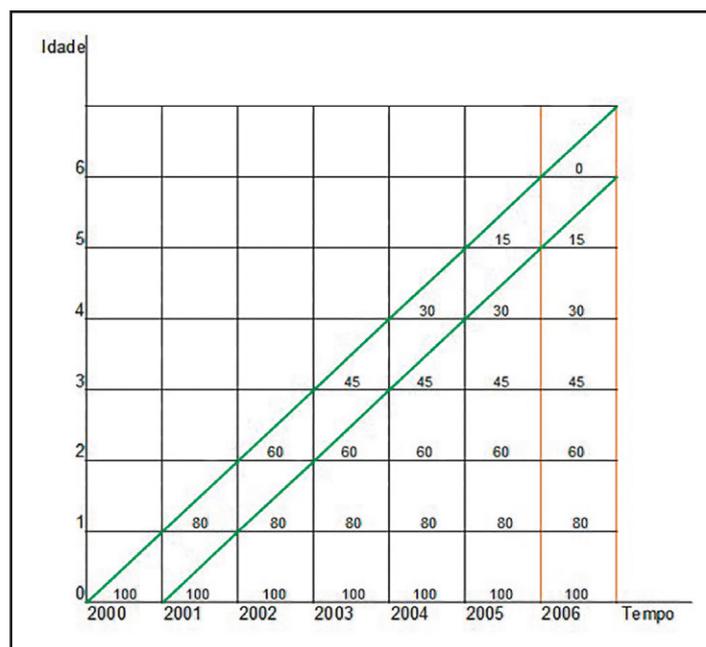
As populações estacionárias possuem uma série de propriedades interessantes do ponto de vista da metodologia demográfica:

- O número de óbitos por idade é sempre o mesmo e o número total de óbitos é igual ao número de nascimentos, que tampouco varia no tempo.
- O tamanho total e a distribuição da população por idades são constantes.
- A estrutura etária da população reproduz exatamente a função  ${}_nL_x$  da tábua de vida, sendo que o número anual de nascimentos equivale à raiz  $\ell_0$  desta tábua.
- Consequentemente, o tamanho da população (por sexo) dividido pelo número anual de nascimentos é igual à esperança de vida (por sexo).
- A TBM e a TBN, ambas são iguais a  $1/e_0$ .

O Capítulo 22 tratará das diferentes relações de uma maneira mais formal, no contexto mais amplo da população estável. Mas para efeitos do capítulo atual, o que importa é particularmente o terceiro ponto, da correspondência entre a estrutura etária de uma população estacionária e sua tábua de vida.

O Gráfico 9.9 mostra a representação de uma população estacionária (dados fictícios) mediante o uso de um Diagrama de Lexis. Pelos dados expostos no diagrama do Gráfico 9.9 nota-se que a composição etária de cada coorte é igual à composição etária da população em cada ano, de forma que a cada ano se registram 100 nascimentos e 100 óbitos e o tamanho da população a cada idade é invariante no tempo. Sob tais circunstâncias é fácil observar que a taxa de crescimento populacional é zero em todas as idades.

Gráfico 9.9: Representação de uma população estacionária no diagrama de Lexis



A situação representada no diagrama do Gráfico 9.9 poderia ser aproximadamente observada numa população real sem migração (população fechada) e as taxas de mortalidade e fecundidade por idade permanecessem constantes por um longo período de tempo.

Em uma população estacionária a composição etária da população reproduz a composição de uma coorte hipotética observada numa tábua de vida. Por essa característica, pode-se fazer relações entre as funções de uma tábua de vida e a população estacionária. Nessa relação, o que numa tábua de vida representa número de anos-pessoa vividos, na população estacionária equivale a número de pessoas efetivamente vivas na população; e enquanto na tábua de vida se refere a pessoas sobreviventes da coorte hipotética a uma exata idade, na população estacionária equivale a pessoas efetivamente vivas a cada tempo com aquela idade exata.

Quadro 9.1: Interpretação das funções da tábua de vida e o seu equivalente na população estacionária

Função	Coorte	População Estacionária
$\ell_x$	Sobreviventes à idade exata $x$	Densidade da população na idade exata $x$
${}_n d_x$	Óbitos entre idades $x$ e $x+n$ , provenientes de uma coorte inicial de $\ell_0$ pessoas	Óbitos anuais que ocorrem entre as idades de $x$ a $x+n$ anos
${}_n a_x$	Tempo médio vivido entre $x$ e $x+n$ por aqueles que morrem neste intervalo	Idade média da população entre $x$ e $x+n$ anos, menos $x$
${}_n L_x$	Anos-pessoa vividos entre $x$ e $x+n$ por uma coorte inicial de $\ell_0$ pessoas	Número de pessoas entre $x$ e $x+n$ anos na população
$T_x$	Anos-pessoa vividos a partir da idade $x$ por uma coorte inicial de $\ell_0$ pessoas	Número de pessoas com $x$ anos ou mais anos na população
$e_x$	Número médio de anos vividos pelos sobreviventes a partir da idade $x$	Idade média da população com $x$ ou mais anos, menos $x$

É importante frisar que certas interpretações que se aplicam a tábuas de vida de coorte perdem a validade no caso de tábuas de período. Por exemplo, numa tábua de coorte, a esperança de vida corresponde à idade média de morte das pessoas que compõem a coorte (ver seção 7.7 do Capítulo 7). A tabela acima esclarece que a esperança de vida na idade  $x$  também pode ser interpretada como o número médio de anos vividos pelos sobreviventes da coorte a partir da idade  $x$  ou como a idade média da população com  $x$  anos ou mais, menos  $x$ , na população estacionária. Entretanto, numa população real, não estacionária, não é correto identificar a esperança de vida com a idade média  $x_D$  dos óbitos. Isso acontece porque  $x_D$ , contrariamente à esperança de vida, depende da estrutura etária da população. Embora  $x_D$  possa ser calculada a partir dos dados sobre óbitos, isso raramente é feito, e a medida que prevalece na análise demográfica é a esperança de vida. Entretanto, a média  $x_D$  tem algumas utilidades na análise das populações estáveis (ver Capítulo 22).

## 9.4 A TÁBUA DE VIDA DE PERÍODO

As tábuas de vida de coorte são fáceis de entender, mas têm dois inconvenientes. Por um lado, elas precisam de séries históricas de dados muito longas. Por outro, elas têm pouco interesse do ponto de vista da descrição das condições de mortalidade na atualidade. Elas respondem perguntas do tipo: “Quantos anos viveram em média as pessoas nascidas há 100 anos ?” Mas não respondem uma pergunta como: “Pelos atuais condições de mortalidade, qual é proporção de pessoas de 15 anos que sobreviverão até os 65 ?” Para responder tais perguntas, o instrumento indicado é a *tábua de vida de período*, também chamada *transversal*. Castro (2015) usa a terminologia *tábuas do segundo tipo*, em oposição às *do primeiro tipo* discutidas acima. Em realidade, a grande maioria das tábuas de vida que são construídas é de período.

Ao final da seção 7.4 foram descritas três maneiras distintas de se quantificar a intensidade dos processos demográficos. Ao construir tábuas de vida de período, considera-se a primeira das três maneiras apresentadas, ou seja, quantificando os eventos demográficos num determinado período para um determinado grupo etário, tendo em mente que tais eventos pertencem a duas coortes diferentes. O recurso metodológico usado para este fim é a *coorte sintética* que foi descrita na seção 7.7.

Para o que segue nesta seção, dois entendimentos se destacam por sua importância. Por um lado, todas as definições, fundamentos e notações feitas para as funções de uma tábua de vida de coorte, na seção anterior, continuam validadas ao trabalhar com dados de período. As diferenças ocorrem por conta das interpretações e formas de mensuração de algumas dessas funções, em decorrência dos dados disponíveis. O segundo entendimento é que, ao estimar uma tábua de vida de período, a interpretação dada às funções está condicionada ao suposto de que a intensidade dos eventos demográficos a cada idade será mantida constante ao longo do tempo, mais especificamente ao longo de uma geração (ou seja, tempo suficiente para que todas as distintas coortes que passam por aquele período sejam extintas). A ideia é de construir uma tábua de vida para caracterizar uma população estacionária a partir das Taxas Específicas de Mortalidade por idade observadas na população atual. Ou seja, ao afirmar que “a esperança de vida de uma criança nascida no ano 2010 no Brasil foi 73,5 anos”, realmente o que está sendo dito é “se a partir de 2010 não houvesse nenhuma mudança nas condições de mortalidade no país, de modo que todas as TEMs permanecessem constantes, uma pessoa nascida em 2010 viveria na média 73,5 anos”. A

vida média real das pessoas nascidas em 2010 quase certamente será diferente, pois as condições de mortalidade mudarão ao longo das suas vidas.

Diferentemente da tábua de coorte que (pelo menos na teoria, embora nem sempre na prática) se baseia na função de sobrevivência  $\ell_x$ , a tábua de período parte da taxa central de mortalidade observada para a população num determinado período de tempo, tal como apresentada em (8.5), geralmente sem o fator de escala 1000:

$${}_nM_{x,s} = \frac{\text{Número de óbitos com idades entre } x \text{ e } x+n \text{ do sexo } s \text{ ocorridos no período}}{\text{População média com idades entre } x \text{ e } x+n \text{ do sexo } s \text{ no período}} = \frac{{}_nD_{x,s}}{{}_n\bar{P}_{x,s}} \quad (9.16)$$

O cálculo da maioria das funções da tábua de vida de período é igual ao cálculo feito na seção 5.2, para a tábua de vida de coorte. A única diferença significativa é que, para calcular  $\ell_x$ , é preciso primeiro estimar a série  ${}_nq_x$  a partir de  ${}_nM_x$ .

Na seção 7.6 argumentou-se sobre a necessidade de se converter as taxas centrais de mortalidade em probabilidades de morte (taxas iniciais do tipo  $q$ ). A taxa central de mortalidade observada (9.16) é uma medida de risco com dados de período, enquanto que a probabilidade  ${}_nq_x$ , definida, por (9.3) é uma medida do risco de morte de uma coorte entre duas idades exatas<sup>7</sup>. Pela definição de  ${}_nq_x$ , se toma a razão dos óbitos entre  $x$  e  $x+n$  ( ${}_nd_x$ ) pelas pessoas sobreviventes a idade exata  $x$  ( $\ell_x$ ), ou seja, o numerador é sempre uma parte do denominador. Acontece que os óbitos no numerador de  ${}_nM_x$  são informações de período e são, portanto, decrementos de duas coortes contíguas. Assim sendo, no cálculo de  ${}_nM_x$ , não necessariamente o numerador é uma parte do denominador. Mais importante do que isso é o fato de que a estrutura etária dentro do intervalo  $(x, x+n)$  não necessariamente corresponde àquela que seria observada numa população estacionária com a mesma tábua de vida subjacente. Apesar disso, a maioria dos métodos que serão discutidos a seguir assumirão que  ${}_nM_x = {}_nm_x$ , a taxa correspondente da população estacionária. A única exceção é o segundo método de Keyfitz, no final desta seção.

Em (9.16), considerando que  ${}_nM_x \approx {}_nm_x$ , tem-se a definição de uma taxa central de mortalidade da tábua de vida:

$${}_nM_x \approx {}_nm_x = \frac{{}_nd_x}{{}_nL_x} \quad (9.17)$$

onde o denominador pode ser interpretado como a população estacionária da tábua de vida entre as idades  $x$  e  $x+n$ . Comparando esta taxa com a probabilidade de morte  ${}_nq_x$  dada por (9.3), nota-se que quando o intervalo etário é igual a 1 ( $n=1$ ) a probabilidade  ${}_nq_x$  é ligeiramente inferior a  ${}_nm_x$ . Já quando  $n=5$ , a  ${}_nq_x$  pode ser cerca de cinco vezes  ${}_nm_x$ . Chama atenção o fato de que ambas as medidas de risco de morte dependem de  $\ell_x$ , que não tem uma formulação matemática explícita e pode variar de uma população para outra. Qualquer suposição sobre a variação de  $\ell_x$  por idade guarda relação direta com a suposição de como os óbitos estão distribuídos ao longo de um período entre duas idades exatas (ver relações 9.6 e 9.8). Isso recoloca o problema encontrado na estimação

<sup>7</sup> Para uma explicação mais detalhada das diferenças entre a taxa central e a probabilidade de morte, ver Capítulo 7.

de  ${}_nL_x$ , ou seja, determinar o tempo médio  ${}_na_x$  vivido entre  $x$  e  $x+n$  por cada pessoa que morreu neste mesmo intervalo. Assim, a conversão de  ${}_nM_x$  (ou, estritamente falando,  ${}_nm_x$ ) em  ${}_nq_x$  passa pelo problema de determinar  ${}_na_x$  ou assumir uma expressão analítica para  ${}_nM_x$ .

A literatura sobre como sobre como converter  ${}_nM_x$  em  ${}_nq_x$  é extensa e a descrição feita aqui será resumida e não exaustiva, com uma referência que aborda cada proposta de forma pormenorizada. As formas de determinação  ${}_nq_x$  com base em  ${}_nM_x$  podem ser divididas em 3 grandes grupos:

1. Determinando  ${}_na_x$  pela observação direta da idade exata ao óbito;
2. Supondo homogeneidade na distribuição dos óbitos ou na intensidade da mortalidade entre duas idades exatas; ou
3. Supondo uma expressão analítica para a função  ${}_nm_x$ .

Determinando  ${}_na_x$  pela observação direta da idade exata ao óbito.

Se os dados de óbitos estão disponíveis na idade exata é possível calcular  ${}_na_x$  diretamente. O problema é que isso requer uma base de dados com informação sobre cada óbito. Também requer que todas as idades das pessoas que morreram sejam conhecidas exatamente, o qual normalmente não é possível sem um mecanismo para relacionar o atestado de óbito com a certidão de nascimento.

Conhecendo  ${}_na_x$  em cada idade e supondo que  ${}_nm_x = {}_nM_x$ , é possível fazer a conversão de  ${}_nM_x$  em  ${}_nq_x$ . A fórmula (9.9) pode ser reescrita como

$${}_nL_x = n \ell_x - (n - {}_na_x) {}_nd_x \quad (9.18)$$

Rearranjando esta expressão, obtém-se

$$\ell_x = \frac{1}{n} ({}_nL_x + (n - {}_na_x) {}_nd_x) \quad (9.19)$$

Agora vale lembrar (9.3) para obter

$${}_nq_x = \frac{{}_nd_x}{\ell_x} = \frac{{}_nd_x}{{}_nL_x + (n - {}_na_x) {}_nd_x} = \frac{{}_nm_x}{1 + (n - {}_na_x) {}_nm_x} \quad (9.20)$$

Embora, estritamente falando, esta relação só vale para  ${}_nm_x$ , supõe-se que ela vale também, com um grau de precisão razoável, para  ${}_nM_x$ . A relação (9.20) possui aplicabilidade, ainda, nos casos em que  ${}_na_x$  é conhecido em outras tábuas de mortalidade, que podem ser consideradas como modelo (padrão) para a população de interesse (Ortega, 1987; Preston, Heuveline e Guillot, 2001).

Supondo homogeneidade na distribuição dos óbitos ou de  ${}_nM_x$  entre duas idades exatas

Se os óbitos forem distribuídos homogeneamente ao longo do intervalo  $(x, x+n)$ ,  $\ell_x$  decresce linearmente e, como já foi visto na seção 7.6,  ${}_na_x = n/2$ . Neste caso, a fórmula (9.20) se reduz para

$${}_nq_x = \frac{{}_nm_x}{1 + 0,5 {}_nm_x} = \frac{2 {}_nm_x}{2 + {}_nm_x} \approx \frac{2 {}_nM_x}{2 + {}_nM_x} \quad (9.21)$$

Esta é mesma fórmula que já foi derivada, por um caminho um pouco diferente, em (7.11).

Por outro lado, supondo que  ${}_nM_x$  é homogêneo, ou seja que a intensidade da mortalidade é igual ao longo do intervalo  $(x, x+n)$ ,  ${}_nq_x$  pode ser escrita como

$${}_nq_x = 1 - \exp(- {}_nM_x) \quad (9.22)$$

A lógica atrás desta fórmula é relativamente simples. Se a intensidade da mortalidade ao longo do intervalo  $(x, x+n)$  é homogênea, a probabilidade de sobrevivência ao longo de uma fração pequena de  $n$ , ou seja  $n/N$ , seria

$${}_{n/N}p_x = 1 - {}_nM_x / N \quad (9.23)$$

Portanto, a probabilidade de sobrevivência para o intervalo inteiro seria

$${}_np_x = (1 - {}_nM_x / N)^N \quad (9.24)$$

Segundo uma regra bem conhecida do cálculo diferencial e integral, para valores grandes de  $N$  (9.24) se aproxima a  $\exp(- {}_nM_x)$ <sup>8</sup>.

As fórmulas (9.21), (9.22) ou eventualmente (9.27) são as mais usadas para converter  ${}_nM_x$  em  ${}_nq_x$ , mas como se verá nos próximos parágrafos, existem várias alternativas mais refinadas.

Supondo uma Expressão Analítica para as Funções  ${}_nm_x$ ,  ${}_nq_x$  ou  $\ell_x$

Algumas das expressões analíticas para o perfil etário das funções  ${}_nm_x$ ,  ${}_nq_x$  e  $\ell_x$  são estabelecidas num intervalo contínuo de idade. Ao analisar  ${}_nm_x$  num intervalo contínuo de idade, obtém-se uma *taxa instantânea de mortalidade* ou *força de mortalidade* a cada instante de idade, a função  $\mu(x)$  que será descrita no Capítulo 17. Por agora, basta saber que a definição matemática da força de mortalidade por idade se constituiu como base teórica para o desenvolvimento de expressões analíticas para as funções  ${}_nm_x$ ,  ${}_nq_x$  e  $\ell_x$ . Na literatura atuarial existem diversas propostas de funções analíticas para descrever  $\mu(x)$  ou  $\ell_x$  que podem ser usadas na conversão de  ${}_nM_x$  ou  ${}_nm_x$  em  ${}_nq_x$ . Para uma revisão excelente destas funções em português se refere aqui à tese de Doutorado de Bravo (2007). Já foi mencionada a Lei de De Moivre ( $\ell_x = 1 - x/\omega$ ), que tem um interesse mais histórico e didático do que prático na construção de tábuas de vida contemporâneas. Algumas outras funções, mais realistas, serão discutidas no Capítulo 20.

<sup>8</sup>  $\exp(x)$  é simplesmente uma forma alternativa para escrever  $e^x$ , com a qual usuários de EXCEL devem estar familiarizados.

### 9.4.1 O método de Reed-Merrell

Entre as propostas de melhoria da relação (9.22), uma das mais conhecidas é o método proposto por Reed e Merrell (1939), que propuseram a seguinte relação para conversão de  ${}_n m_x$  em  ${}_x q_x$ :

$${}_n q_x = 1 - \exp(-n {}_n m_x - a \cdot n^3 \cdot {}_n m_x^2) \quad (9.25)$$

onde  $a = 0,008$ . Como sempre, na prática a fórmula é aplicada com  ${}_n M_x$ , em vez de  ${}_n m_x$ . Pressat (1967) argumenta sobre uma aplicabilidade limitada da proposta de Reed-Merrell, especialmente nas primeiras idades (0-1 e 1-4 anos). O método Reed-Merrell foi desenvolvido com base na experiência de mortalidade observada em diferentes regiões dos EUA entre 1910 e 1930 e pode não refletir diferentes situações de diversas populações. Entretanto, para grupos etários acima dos 5 anos a proposta de Reed-Merrell ainda é de grande interesse (Pressat, 1967). Ademais, o método permite incorporar, em sua forma analítica, fatores de correção para omissão de óbitos nas primeiras idades (Ortega, 1987: 134). Uma demonstração da relação (9.25) pode ser encontrada em Ortega (1987) ou em Reed e Merrell (1939).

### 9.4.2 O método de Greville

Um método proposto por Greville (1943) tem grande aplicabilidade quando o tamanho do intervalo etário  $n$  varia de um grupo etário para outro. A relação proposta por Greville (1943) para converter  ${}_n M_x$  em  ${}_n q_x$  é:

$${}_n q_x = \frac{{}_n M_x}{\frac{1}{n} + {}_n M_x \left[ \frac{1}{2} + \frac{n}{12} ({}_n M_x - \ln(c)) \right]} \quad (9.26)$$

onde  $c$  é uma constante que varia entre 0,080 e 0,104 vem do pressuposto que  ${}_n M_x$  segue a Lei de Gompertz (ver Capítulo 20). Segundo Ortega (1987) o valor de  $c$  tem pouco efeito para a estimação de  ${}_n q_x$  nas idades iniciais uma vez que a Lei de Gompertz é mais forte nas idades avançadas. Para  ${}_n q_x$  abaixo dos 5 anos, intervalo etário em que os eventos vitais são grandemente afetados por omissões, Ortega (1987) sugere a aplicação de outros procedimentos.

### 9.4.3 Procedimento iterativo

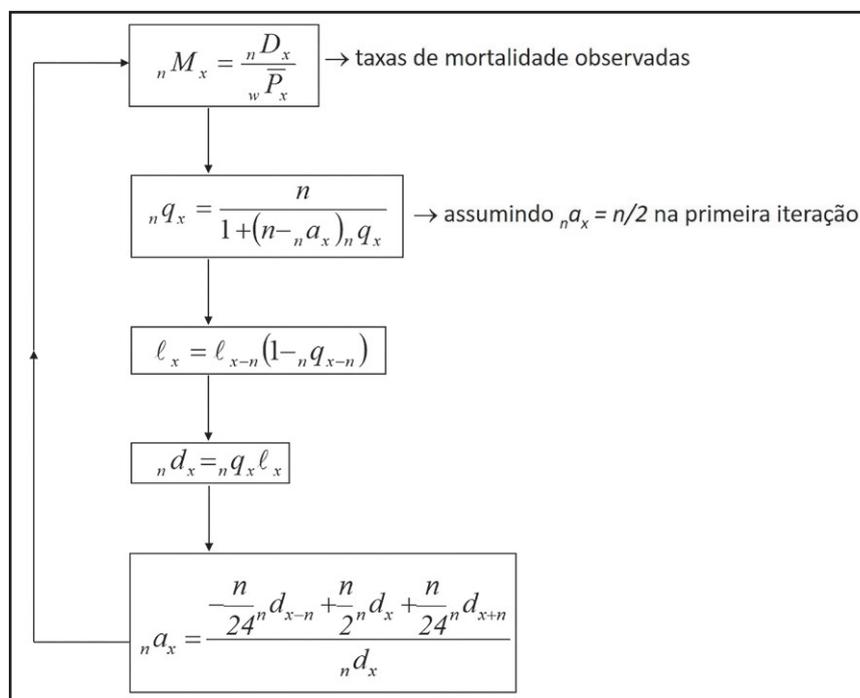
Keyfitz (1966) propôs um procedimento iterativo para estimar  ${}_n q_x$  mediante  ${}_n M_x$  assumindo que a distribuição de óbitos na tábua de vida ( ${}_n d_x$ ) segue uma função polinomial de segundo grau no intervalo  $(x - n, x + 2n)$ . Sob essa suposição,  ${}_n a_x$  pode ser estimada por:

$${}_n a_x = \frac{-\frac{n}{24} {}_n d_{x-n} + \frac{n}{2} {}_n d_x + \frac{n}{24} {}_n d_{x+n}}{{}_n d_x} \quad (9.27)$$

A prova desta afirmação não é difícil, mas ela não é mostrada aqui porque ela envolve a integração das curvas  $A \cdot x^2 + B \cdot x + C$  e  $A \cdot x^3 + B \cdot x^2 + C \cdot x$  nos intervalos  $(x-n, 0)$ ,  $(0, x+n)$  e  $(x+n, x+2n)$ . Ela sim é mostrada na seção 17.3 do Capítulo 17.

Se  ${}_n d_x$  está simetricamente distribuído ao longo dos três intervalos etários em (9.27),  ${}_n a_x$  se reduz a  $n/2$ . O problema na relação (9.27) é que  ${}_n a_x$  depende de  ${}_n d_x$ , que por sua vez depende de  ${}_n q_x$  (por estimar). A proposta de Keyfitz (1966) é, então, iniciar um procedimento iterativo assumindo inicialmente que  ${}_n a_x = n/2$ . Em seguida, aplica-se a relação (9.27) para obter uma primeira estimativa de  ${}_n q_x$  e, conseqüente-mente,  ${}_n d_x$ . Então, utiliza-se a estimativa preliminar de  ${}_n d_x$  em (9.27) para se obter uma segunda estimativa de  ${}_n a_x$ . Esse ciclo iterativo é mostrado na Figura 9.1 e se repete até que os valores de  ${}_n a_x$  se estabilizem ao redor da quarta casa decimal. Antes de aplicar o procedimento propriamente dito, é recomendável suavizar a sequência dos  ${}_n M_x$ . Embora seja possível executar o processo todo em EXCEL, é mais conveniente usar “R”, tanto para facilitar a suavização como para lidar com o processo iterativo, que é mais fácil de implementar em “R” do que em EXCEL. Um programa desenvolvido para esta finalidade é mostrado como exemplo do uso de “R” na primeira seção do Capítulo 17.

Figura 9.1: Método de iteração para reproduzir as taxas específicas de mortalidade na população



Fonte: Keyfitz (1966) e Preston, Heuveline e Guillot (2001).

As formas alternativas para derivar  ${}_n q_x$  mediante  ${}_n M_x$  possuem muita aplicabilidade em Demografia e Ciências Atuariais. A escolha por determinado procedimento deve ser pautada pelo tipo de tábua de vida que se deseja: se abreviada ou completa. Ademais, a combinação entre procedimentos que são mais apropriados para determinado intervalo de idade é uma boa alternativa.

#### 9.4.4 A segunda proposta de Keyfitz

Além da proposta iterativa descrita acima, Keyfitz propôs um segundo método que, como o método de Reed-Merrell, é um refinamento de (9.22) (Keyfitz e Frauenthal, 1975; Keyfitz, 1977). Este método é digno de nota por ser o único que não assume que  ${}_n m_x = {}_n M_x$  e que explicitamente considera a estrutura etária da população subjacente. A fórmula proposta é a seguinte:

$${}_n q_x = 1 - \exp\{-n {}_n M_x + n ({}_n P_{x+n} - {}_n P_{x-n}) ({}_n M_{x+n} - {}_n M_{x-n})\} / 48 {}_n P_x \quad (9.28)$$

Segundo Pressat (1995), esta fórmula funciona bem em situações onde a estrutura etária da população é relativamente suave, mas em situações onde há variações bruscas de uma faixa etária para outra, ela não necessariamente dá melhores resultados do que outros métodos mais simples. Por isso, é recomendável graduar (suavizar) os dados antes de aplicar (9.28). Na seção 17.3 do Capítulo 17 será proposto um procedimento alternativo, implementado por meio de um programa em “R”, que usa uma interpolação quadrática, parecida com a usada em (9.27), em vez da linear usada em (9.28), para derivar  ${}_n q_x$  de  ${}_n M_x$  com a consideração da estrutura etária da população.

#### 9.4.5 ${}_n a_x$ para tábuas de vida abreviadas

As tábuas de vida são construídas, geralmente, por grupos etários quinquenais, exceto nas primeiras idades, onde não se pode assumir uniformidade no tempo médio vivido pelos que morreram. Ou seja, quando uma população inicia sua transição da mortalidade, a queda nas taxas de mortalidade é mais acentuada no primeiro ano de vida. Este pressuposto de não uniformidade é ainda mais evidente no primeiro ano de vida, onde os determinantes para reduções nas taxas de mortalidade infantil são classificados em exógenos e endógenos que atuam com maior intensidade, respectivamente, nas componentes, neonatal (desde o nascimento até o vigésimo sétimo dia) e pós-neonatal (desde o vigésimo oitavo dia até penúltimo dia para completar um ano exato de vida). Os fatores exógenos que determinam quedas na mortalidade pós-neonatal estão relacionados, em sua maioria, a condições ambientais e imunológicas desfavoráveis, como falta de saneamento básico e de acesso a vacinas. Então, no início da transição da mortalidade, esses seriam os fatores mais preponderantes para a redução do risco de morte infantil, fazendo com que os óbitos infantis se concentrem nas primeiras semanas de vida. Já os fatores endógenos, que atuam na queda da mortalidade infantil neonatal, estão relacionados, em grande medida, aos avanços tecnológicos na medicina e aos cuidados básicos durante o período de pré-natal. A atuação destes fatores na redução da componente neonatal da mortalidade infantil relaciona-se positivamente com o nível de desenvolvimento econômico, social e educacional da população. Assim, tais componentes de redução da mortalidade infantil e a forma como atuam seus respectivos determinantes fazem com que os óbitos no primeiro ano de vida se concentrem nas primeiras semanas.

No caso de tábuas de vida abreviadas, se os primeiros grupos etários são 0-1 e 1-4 anos e os dados disponíveis não possibilitam calcular  ${}_n a_x$  diretamente, alguns procedimentos estão disponíveis. Primeiramente, destaca-se uma proposta de Bourgeois-Pichat (1951 a b) para representação analítica da mortalidade no primeiro ano de vida (ver também Pressat: 453). Coale e Demeny (1983: 20) sugerem o uso de valores de  ${}_n a_x$  determinados empiricamente em função do

nível de mortalidade traduzido pela magnitude de  ${}_1M_0$  (taxa observada de mortalidade abaixo de 1 ano de idade). Os autores analisaram a relação entre  ${}_1a_0$  e  ${}_4a_1$  com valores de  ${}_1q_0$  e ajustaram coeficientes a serem aplicados a  ${}_1M_0$  em função dos próprios valores de  ${}_1M_0$  proveniente dos dados observados. Preston, Heuveline e Guillot (2001) sintetizaram a aplicação desses coeficientes na tabela abaixo:

Tabela 9.2: Coeficientes para determinação dos valores de  ${}_na_x$  abaixo dos 5 anos de idade

	Homens	Mulheres
<b>Valores de <math>{}_1a_0</math></b>		
Se ${}_1m_0 \geq 0,107$	0,330	0,350
Se ${}_1m_0 < 0,107$	$0,045 + 2,684 \cdot {}_1m_0$	$0,053 + 2,8 \cdot {}_1m_0$
<b>Valores de <math>{}_4a_1</math></b>		
Se ${}_1m_0 \geq 0,107$	1,352	1,361
Se ${}_1m_0 < 0,107$	$1,651 - 2,816 \cdot {}_1m_0$	$1,522 - 1,518 \cdot {}_1m_0$

Nota:  ${}_1m_0$  é a taxa específica de mortalidade de 0-1 ano para ambos os sexos. Não é a taxa de mortalidade infantil.

Fonte: Preston, Heuveline e Guillot (2001): Table 3.3.

As tábuas de vida preparadas pelo IBGE em 2013 usam fatores de  ${}_1a_0 = 0,089291033$  para homens e  $0,0910919718$  para mulheres, que são valores muito próximos aos recomendados na Tabela 9.2.

#### 9.4.6 Finalizar a tábua de vida

O outro grupo etário que requer um tratamento especial é o intervalo aberto final. Sabe-se que a probabilidade condicional de morte neste grupo de idade é sempre igual 1, pois, dado que os indivíduos chegaram vivos à idade inicial do intervalo aberto (80 anos, por exemplo), a probabilidade de morte acima dessa idade é igual a 1 pois todos morrerão neste intervalo. Sabe-se, ainda, que o tempo médio a ser vivido pelos que morreram neste intervalo é inversamente proporcional à taxa de mortalidade neste grupo etário. Portanto, o valor de  ${}_na_x$  no último grupo de idade é, por definição, a esperança de vida neste grupo etário:

$${}_{\omega}a_x = e_x \quad (9.29)$$

Um resultado que deriva de (9.29) é que, se  ${}_{\omega}a_x$  é igual à esperança de vida do último grupo etário, é possível escrever  ${}_{\omega}a_x$  em função de  ${}_{\omega}m_x$  a partir da definição de  $e_x$ , dada por (9.14):

$${}_{\omega}a_x = e_x = \frac{T_x}{l_x} = \frac{{}_{\omega}L_x}{{}_{\omega}d_x} = \frac{1}{{}_{\omega}m_x} \quad (9.30)$$

Outra forma de se estimar a esperança de vida no último grupo etário quando não se confia na informação observada de  ${}_{\omega}M_x$  é estimar  ${}_nL_x$  no último grupo etário ( ${}_{\omega}L_x$ ). O cálculo de  ${}_{\omega}L_x$  depende da longevidade da coorte, ou seja, seria necessário ter a informação da coorte até a sua completa extinção. Quando não se dispõe dessa informação, o cálculo de  ${}_{\omega}L_x$  depende da taxa central de mortalidade derivada de uma tábua de vida descrita em (9.17). Da relação (9.17), quando  $n$  se refere ao intervalo etário aberto, obtém-se:

$${}_{\omega}L_x = \frac{{}_{\omega}d_x}{{}_{\omega}m_x} \quad (9.31)$$

Entretanto, o número de óbitos no último intervalo  ${}_{\omega}d_x$  etário é igual ao número de sobreviventes que alcançam este grupo de idade  $\ell_x$ . Portanto:

$${}_{\omega}L_x = \frac{\ell_x}{{}_{\omega}m_x} \approx \frac{\ell_x}{{}_{\omega}M_x} \quad (9.32)$$

onde, como antes, se assume que  ${}_nM_x$  é igual a  ${}_nm_x$  (taxa de mortalidade da tábua, ou seja, da população estacionária).

Os problemas potenciais na aplicação de (9.32) são que  ${}_nM_x$  pode ser substancialmente diferente de  ${}_nm_x$ , que os números pequenos de óbitos observados podem causar muita variação aleatória em  ${}_nM_x$  e que os resultados podem ser distorcidos por erros na declaração das idades. Por isso muitas vezes se adotam procedimentos alternativos. O INE (2007) de Portugal usa um procedimento log-linear proposto por Denuit e Goderniaux (2005) para ajustar uma curva teórica aos últimos valores observados de  ${}_nq_x$  e extrapolar a curva até os 95 anos. Outros países mais desenvolvidos adotam procedimentos parecidos de extrapolação. A seção 9.6 mostra um procedimento que tem sido usado pelo IBGE no Brasil.

Uma alternativa é usar o último valor de  $\ell_x$  ou de  ${}_nq_x$  para extrapolar a tábua de vida, com base numa tábua de vida modelo (ver seção 20.2.3 do Capítulo 20). Também existem fórmulas abreviadas para este fim. Por exemplo, Ortega (1987: 33) sugere as seguintes fórmulas para  ${}_{\omega}L_x$  ( $T_x$ ) e conseqüentemente para  $e_x$ :

$${}_{\omega}L_{75} = (5,731 + 0,0000654 \ell_{75}) \ell_{75} \quad (9.33.a)$$

$${}_{\omega}L_{80} = (4,769 + 0,0000536 \ell_{80}) \ell_{80} \quad (9.33.b)$$

$${}_{\omega}L_{85} = (3,862 + 0,0000466 \ell_{85}) \ell_{85} \quad (9.33.c)$$

Coale e Demeny (1966) usaram  ${}_{\omega}L_{85} = \ell_{85} \log_{10}(\ell_{85})$ . Entretanto, todas estas fórmulas mais antigas tendem a sobre-estimar a mortalidade nas idades mais avançadas. A seguinte fórmula funciona bem para reproduzir os valores de  $e_{85}$  usados pela Divisão de População das Nações Unidas nas suas estimativas e projeções posteriores a 2010, para tábuas com  $50 \leq e_0 \leq 85$ :

$$e_{85} = 7,282 (-\ln({}_5q_{80} + 0,0943))^{0,796} \quad (9.34)$$

#### 9.4.7 Exemplo de uma tábua de coorte

A Tabela 9.3 dá uma ideia de como funciona a mecânica da construção de uma tábua de vida elaborada a partir dos valores de  ${}_n m_x$  da Tabela 8.3. Trata-se de uma planilha que pode ser montada facilmente em EXCEL. Se a primeira linha for 1 e a primeira coluna A, se a Coluna A contiver os valores de  $x$  e a coluna B os de  $n$ , e se a coluna C contiver os valores dados de  ${}_n m_x$  e a coluna D os multiplicadores  ${}_n a_x$ , as demais colunas podem ser obtidas da seguinte forma:

${}_+ a_{80}$ :	D19 = 1/C19
${}_l q_0$ até ${}_+ q_{75}$ :	E2= (B2*C2)/(1+(B2-D2)*C2) e assim para baixo até a linha 19
${}_+ q_{80}$ :	E19=1
$\ell_0$ :	F2=100000
$\ell_1$ até $\ell_{80}$ :	F3 = F2*(1-E2) e assim para baixo até a linha 19
${}_l d_0$ até ${}_+ d_{80}$ :	G2 = F2 – F3 e assim para baixo até a linha 19
${}_l L_0$ até ${}_+ L_{80}$ :	H2= F3*B2 + G2*D2 e assim para baixo até a linha 19 (com F20 = 0)
$T_{80}$ :	I19 = H19
$T_0$ até $T_{75}$ :	I18 =I19+H18 e assim para cima até a linha 2
$e_0$ até $e_{80}$ :	J2 = I2/F2 e assim para baixo até a linha 19

Os valores de  ${}_l a_0$  e  ${}_+ a_1$  na Tabela 9.3 foram obtidos utilizando a Tabela 9.2. Já o valor de  ${}_{\omega} a_{80}$  foi obtido pela fórmula (9.30).

Tabela 9.3: Tábua de vida masculina, Brasil, 2016

	A	B	C	D	E	F	G	H	I	J
1	x	N	${}_n m_x$	${}_n a_x$	${}_n q_x$	$l_x$	${}_n d_x$	${}_n L_x$	$T_x$	ex
2	0	1	0,014362	0,08	0,0142	100000	1418	98701	7184094	71,84
3	1	4	0,000535	1,61	0,0021	98582	211	393826	7085393	71,87
4	5	5	0,000240	2,5	0,0012	98372	118	491564	6691567	68,02
5	10	5	0,000365	2,5	0,0018	98254	179	490821	6200003	63,10
6	15	5	0,002061	2,5	0,0103	98075	1005	487859	5709182	58,21
7	20	5	0,002701	2,5	0,0134	97069	1302	482090	5221322	53,79
8	25	5	0,002534	2,5	0,0126	95767	1206	475821	4739232	49,49
9	30	5	0,002571	2,5	0,0128	94561	1208	469787	4263411	45,09
10	35	5	0,003030	2,5	0,0150	93353	1404	463258	3793624	40,64
11	40	5	0,003855	2,5	0,0191	91950	1755	455360	3330366	36,22
12	45	5	0,005376	2,5	0,0265	90194	2392	444991	2875005	31,88
13	50	5	0,007687	2,5	0,0377	87802	3311	430733	2430014	27,68
14	55	5	0,011333	2,5	0,0551	84491	4656	410816	1999281	23,66
15	60	5	0,016569	2,5	0,0795	79835	6351	383299	1588465	19,90
16	65	5	0,024469	2,5	0,1153	73484	8472	346242	1205166	16,40
17	70	5	0,037869	2,5	0,1730	65012	11245	296948	858924	13,21
18	75	5	0,057972	2,5	0,2532	53767	13612	234805	561976	10,45
19	80	+	0,122734	8,15	1,0000	40155	40155	327171	327171	8,15

Fonte: Tabela 8.3.

### 9.4.8 Graduação

Graduação se refere à modificação dos dados originais para obter uma série mais suave ou regular (mais consistente com o perfil teórico esperado) antes de aplicar os métodos de estimação explicados nas seções 9.2 e 9.4. A graduação pode ser aplicada a dados de mortalidade, fecundidade, migração ou estrutura etária da população, embora o seu uso no ajuste de dados de mortalidade historicamente tenha sido o mais comum. Os dois objetivos principais da graduação são:

1. Remover flutuações aleatórias nos dados decorrentes do fato que eles se baseiam em números limitados (e às vezes relativamente pequenos) de observações que podem variar de uma forma errática de um ano para outro ou de uma idade para outra.
2. Remover distorções sistemáticas decorrentes de erros mais ou menos previsíveis, como a tendência a arredondar idades.

Não se trata da correção da informação de base para eliminar erros sistemáticos no nível de cobertura. Para isso existem outros métodos que serão discutidos no Capítulo 23. A graduação só melhora a consistência interna das séries de dados.

Quanto ao primeiro ponto, vale lembrar que as funções da tábua de vida e especificamente a esperança de vida podem de um ano para outro, não devido a alguma mudança estrutural no estado de saúde da população, mas simplesmente como resultado de flutuações aleatórias. Por exemplo, a esperança de vida da população masculina de Cabo Verde em 2010-2014, segundo a Divisão de População das Nações Unidas, era 67,91 anos (69,74 segundo o INE de Cabo Verde). Mas devido ao fato de que esta estimativa se baseia numa população de meros 250 mil pessoas, o intervalo de confiança de 95% se estende de 67,15 anos até 68,60 anos. Os detalhes destes cálculos vão além dos objetivos deste livro. Os textos recomendados sobre este assunto são os de Chiang (1968, 1984). Só como regra de punho pode-se considerar que a amplitude do intervalo de confiança de 95% para uma população de 1 milhão é aproximadamente 0,7 anos e que esta quantidade diminui de forma inversamente proporcional com a raiz quadrada da população. Assim, a esperança de vida numa população de 100.000 pessoas teria um intervalo de confiança de aproximadamente  $0,7 \cdot \sqrt{10}$ , ou seja, 2,20 anos. Isto demonstra que calcular esperanças de vida a nível municipal pode ser problemático para municípios relativamente pequenos, mesmo quando se dispõe de dados detalhados e confiáveis sobre o número de óbitos por idade.

A graduação não pode evitar os problemas descritos no parágrafo anterior, mas sim, pode reduzir a incidência de padrões etários irrealistas devido a flutuações erráticas dos dados por idade. Os métodos que existem para este propósito podem ser divididos em métodos gerais, que podem ser aplicados a qualquer tipo de dados, e métodos especificamente desenvolvidos para dados de mortalidade, fecundidade ou migração. Estes últimos usam algum tipo de informação para modelar o perfil esperado dos fenômenos estudados.

Antes de entrar no processo de graduação propriamente dito, há duas considerações importantes por fazer:

1. As irregularidades observadas nos dados realmente se devem a erros ou poderiam ser variações reais? Quando os dados se referem a números relativamente pequenos de observações e a preocupação principal é com a eliminação da variabilidade aleatória, esta pergunta pode não ser muito premente, mas em situações onde algumas idades ou algumas unidades de tempo apresentam condições especiais, ela merece séria consideração. Por exemplo, um surto de uma doença contagiosa nas escolas poderia causar níveis de mortalidade particularmente elevados nas idades correspondentes. Antes de suavizar estes dados, efetivamente apagando informação real, é importante averiguar se isso se justifica.
2. Qual é a variável que deve ser graduada? No caso de  ${}_nM_x$ , será a própria taxa ou será melhor graduar o número de óbitos e a população subjacente separadamente? Tanto o numerador como o denominador poderiam conter erros e estes erros poderiam se reforçar mutuamente ou poderiam tender a se neutralizar. Por exemplo, se tanto os óbitos como a população subjacente têm a tendência a se concentrar em anos que terminam em “0” e as duas tendências têm intensidades parecidas, o resultado  ${}_nM_x$  poderia ser aproximadamente correto e não precisar de ajustes. Por outro lado, se os padrões de erro do numerador e do denominador forem diferentes, exigindo alguma redistribuição de ambos, é melhor lidar com cada problema separadamente. A redistribuição de óbitos e de população não muda o número total de óbitos ou de pessoas, enquanto uma graduação direta de  ${}_nM_x$  poderia afetar

a relação entre os totais. Uma outra consideração poderia ser a existência de algum modelo para ser aplicado aos dados. Pode haver modelos teóricos para a graduação de  ${}_nM_x$ , mas não para os óbitos e populações subjacentes.

3. O objetivo é obter uma sequência de pontos discretos ajustados, igual ao formato dos dados originais, ou procura-se uma função contínua com características suaves para descrever os dados discretos de base ?

### 9.4.9 Métodos gerais

Vários dos métodos gerais são extremamente simples e intuitivos. Um método muito usado e que já foi mencionado em (8.4) consiste em juntar os dados de anos sucessivos para reduzir a flutuação. Por exemplo, para estimar  ${}_nM_x$  no ano 2000, toma-se a média do número de óbitos em 1999, 2000 e 2001 no numerador e a população estimada para a metade do ano 2000 no denominador de  ${}_nM_x$ . No que diz respeito ao detalhe por idade, normalmente as idades são agrupadas em faixas etárias de 5 anos, com o mesmo objetivo. Se os maiores problemas têm a ver com o arredondamento de idades para números que terminam em “0” ou “5”, pode valer a pena agrupar os dados em intervalos centrados nestas idades e depois usar algum procedimento de interpolação para dividi-los em idades simples. Da mesma forma, alguns autores recomendam que, dependendo da qualidade da informação, em alguns casos pode ser melhor construir tábuas de vida completas pela desagregação de tábuas de vida abreviadas, em vez de fazê-lo diretamente a partir das taxas por idade simples (Kostaki e Panousis, 2001).

Uma alternativa é o uso de *médias móveis*, que fornecem resultados por idade simples, mas baseados na média da idade em questão e as idades vizinhas. As médias móveis também podem ser aplicadas de forma ponderada, dando maior peso para os pontos mais próximos e menos para os mais distantes. Por exemplo, uma média móvel baseada em 5 pontos poderia ter a seguinte ponderação:

$$\overline{M}_x = (M_{x-2} + 2 M_{x-1} + 3 M_x + 2 M_{x+1} + M_{x+2})/9 \quad (9.35)$$

Uma alternativa mais forte é

$$\overline{M}_x = (M_{x-4} + 2 M_{x-3} + 3 M_{x-2} + 4 M_{x-1} + 5 M_x + 4 M_{x+1} + 3 M_{x+2} + 2 M_{x+3} + M_{x+4})/25 \quad (9.36)$$

Os mesmos procedimentos também podem ser aplicados separadamente para óbitos e populações. Um inconveniente das médias móveis é que elas não funcionam no início e no final da série, onde faltam valores vizinhos. Uma outra limitação é que elas reproduzem a mesma estrutura discreta dos dados originais quando em muitos casos pode ser mais interessante definir uma função contínua para descrever os dados.

Uma técnica um pouco mais avançada e flexível é a fórmula de Whittaker-Henderson, geralmente aplicada para suavizar uma série observada de valores  $q_x$  para idades simples. A lógica do método consiste em minimizar uma das seguintes expressões:

$$\sum_x w_x (q_x - \bar{q}_x)^2 + w (\bar{q}_{x+1} - 2\bar{q}_x + \bar{q}_{x-1})^2 \quad (9.37.a)$$

ou

$$\sum_x w_x (q_x - \bar{q}_x)^2 + w (\bar{q}_{x+2} - 3\bar{q}_{x+1} + 3\bar{q}_x - \bar{q}_{x-1})^2 \quad (9.37.b)$$

A primeira parte destas expressões minimiza os desvios entre os valores observados e os suavizados, enquanto a segunda parte assegura que a série de valores suavizados seja regular, sendo que o critério (9.37.a) implica uma suavização mais forte do que (9.37.b). Ao escolher os valores de  $w_x$  e  $w$ , é possível dar mais peso a uma correspondência próxima com os valores observados, ou então a um comportamento mais suave da série ajustada. Como  $w_x$  depende da idade  $x$ , é possível variar o seu valor, dando mais peso a algumas idades do que a outras. Existem diferentes técnicas para calcular a série ajustada. Antigamente, a quantidade de valores por estimar fazia com que, em geral, a solução do sistema era relativamente complexa, exceto no caso onde  $w_x = 1$  para todos os valores de  $x$ , conhecido como a fórmula de Whittaker-Henderson Tipo A. Hoje em dia a forma mais simples para resolver o caso geral é pelo uso do recurso Solver de EXCEL, que permite estimar até 200 variáveis simultâneas. Gomes, Okubo e Vasconcelos (2008) explicam o método em mais detalhe.

Existem outras técnicas que produzem resultados mais satisfatórios neste sentido, que são

- Os métodos LOESS e LOWESS, de regressão local;
- Os métodos de kernel (não serão tratados aqui);
- Os métodos bayesianos de graduação;
- A interpolação osculatória; e
- Os métodos de splines.

Como se trata de métodos matematicamente mais avançados, a sua discussão será adiada até mais tarde. Os métodos LOESS e LOWESS e de kernel serão discutidos no Capítulo 16. Os métodos bayesianos serão brevemente discutidos na seção 17.5 do Capítulo 17 e os métodos de interpolação osculatória serão introduzidos no Capítulo 18.

#### 9.4.10 Métodos específicos

Nas seções anteriores deste capítulo já se fez referência a certos supostos analíticos sobre a forma da curva  $\ell_x$  ou da intensidade da mortalidade. Especificamente foram mencionadas as

curvas de Gompertz, Makeham e Heligman-Pollard. É importante frisar que até agora se tratava da forma destas curvas *dentro* de cada faixa etária e as suas consequências para a conversão de  ${}_n m_x$  em  ${}_n q_x$  ou para o cálculo de  ${}_n a_x$ . No caso da graduação, as curvas analíticas são aplicadas a intervalos etários mais longos para *modificar* os valores observados, para que se conformem ao padrão etário esperado. As curvas de Gompertz e Makeham têm uma longa história nos estudos atuariais, onde foram aplicadas basicamente para descrever a mortalidade adulta. Elas não descrevem adequadamente o perfil etário abaixo dos 30 anos de idade e também há indicações de que elas exageram os níveis de mortalidade acima dos 85 anos. A curva de Heligman e Pollard (1980) é mais complexa, mas se adapta melhor ao padrão de mortalidade em todas as idades. Um outro critério para a graduação de tábuas de vida observadas são os sistemas de tábuas de vida modelo, particularmente os desenvolvidos pela Universidade de Princeton (Coale e Demeny, 1983) e pela Divisão de População das Nações Unidas (Naciones Unidas, 1983). Todos esses modelos serão discutidos em mais detalhe no Capítulo 20.

## 9.5 BASES DE INFORMAÇÃO COM TÁBUAS DE VIDA

Como já foi discutido no Capítulo 5, existem várias bases de dados sobre mortalidade e mais especificamente tábuas de vida. A *Human Mortality Data Base* ([www.mortality.org](http://www.mortality.org)) (Barbieri et al., 2015) já foi mencionada neste capítulo. Em 2015 a base continha tábuas de 45 países, em 11 deles tanto de período como de coorte. A Organização Mundial da Saúde (OMS) também mantém uma base de dados sobre mortalidade ([www.who.int/healthinfo/mortality\\_data/en/](http://www.who.int/healthinfo/mortality_data/en/)) e mais especificamente sobre tábuas de vida (<http://apps.who.int/gho/data/?theme=main&vid=60000>) que contém tábuas de vida para todos os países do mundo. A outra base de dados que fornece tábuas de vida atualizadas para todos os países do mundo é aquela mantida e atualizada de dois em dois anos pela Divisão de População das Nações Unidas, como parte dos dados usados na preparação das suas projeções demográficas (<https://esa.un.org/unpd/wpp/Download/Standard/Mortality/>). Embora ambas as bases sejam mantidas por agências das Nações Unidas, elas não são necessariamente consistentes e as tábuas de vida obtidas de uma ou outra base podem ser substancialmente diferentes.

O Capítulo 5 também fez referência à *Latin American Human Mortality Data Base*, mantida por pesquisadores da Universidade Nacional da Colômbia e da Universidade Federal de Minas Gerais no Brasil ([www.lamortalidad.org](http://www.lamortalidad.org)), com dados da Argentina, Brasil, Colômbia, México e Peru, e da *Latin American Data Base* (LAMBdA) mantida por pesquisadores da Universidade de Wisconsin e da Universidade de Califórnia (UCLA) ([www.ssc.wisc.edu/cdha/latinmortality/](http://www.ssc.wisc.edu/cdha/latinmortality/)). Estas bases cobrem um período histórico bem mais extensor do que a OMS que dá prioridade a dados contemporâneos. A base LAMBdA aplica correções e suavizações extensas aos dados originais, mas em contrapartida é mais restritiva na inclusão de dados. A outra base tem uma cobertura mais ampla, incluindo dados subnacionais e em anos intercensitários, mas representa os dados mais ou menos na sua forma original, só com correções dos problemas mais evidentes. Finalmente há que mencionar as tábuas de vida do Brasil preparadas pelo IBGE, que podem ser encontradas em <https://www.ibge.gov.br/estatisticas-novoportal/sociais/populacao/9126-tabuas-completas-de-mortalidade.html?&t=o-que-e>. As tábuas de vida completas de Portugal também podem ser descarregadas do sítio web do INE.

## 9.6 TÁBUAS DE VIDA ATUARIAIS E POPULACIONAIS

Tábuas de vida são largamente utilizadas em cálculos atuariais, os quais consideram a probabilidade dos pagamentos acontecerem no decorrer de vários anos futuros para determinar o valor necessário para pagar os benefícios em planos previdenciários, seguros de vida, ou outros tipos de seguros, tanto estatais como privados. Um benefício previdenciário, por exemplo, pode durar enquanto o segurado estiver vivo, o que é expresso pela tábua de vida. Já nos seguros de vida o segurado paga prêmios à seguradora enquanto estiver vivo e a seguradora paga aos dependentes do segurado um valor maior quando este vir a falecer. Saber a probabilidade de morte do segurado a cada ano, portanto, significa definir com mais certeza o quanto a seguradora poderia aplicar em investimentos de longo ou de curto prazo, e os valores que os segurados precisariam pagar de forma a, para a população de segurados, acumular-se o valor necessário para o pagamento aos dependentes.

Pela natureza de sua aplicação, as tábuas de vida utilizadas em cálculos atuariais têm algumas particularidades. Nessas tábuas, a menos que se trate de pagamento de pensões a menores, os detalhes da mortalidade até os 18 anos de idade geralmente são pouco relevantes, já que nessas idades a probabilidade de morte é pequena em relação aos demais grupos etários. Por outro lado, as tábuas de vida atuariais precisam fornecer maiores detalhes sobre a população de 65 anos e mais do que acontece com as tábuas de vida gerais. Em tábuas de vida gerais é comum observarmos grupos etários quinquenais e o último grupo etário aberto englobando uma grande quantidade de possibilidades de idade à morte. Nas tábuas utilizadas para cálculos atuariais é essencial que as probabilidades de morte sejam definidas para idades simples, inclusive no último grupo etário, que é o que detém maior probabilidade de morte e, portanto, de pagamentos. Isso se dá porque não interessa saber apenas o tempo médio de duração dos benefícios ou pagamentos, mas a distribuição da probabilidade de morte entre as idades, de forma a se determinar, com mais exatidão, o fluxo dos pagamentos esperados. Assim, é comum tabelas de vida para cálculos atuariais em idades simples que vão de 0 até os 110 ou 115 anos. Considerando essa particularidade, o IBGE divulga anualmente, no site responsável pelas informações previdenciárias oficiais no Brasil, uma tábua de vida por idade simples extrapolada até mais de 100 anos, a ser utilizada nos cálculos atuariais, sobretudo nos referentes à previdência de servidores públicos. Os logaritmos da probabilidade de morte para o sexo masculino por essa tábua para o ano de 2016 para homens são apresentados no Gráfico 9.10.

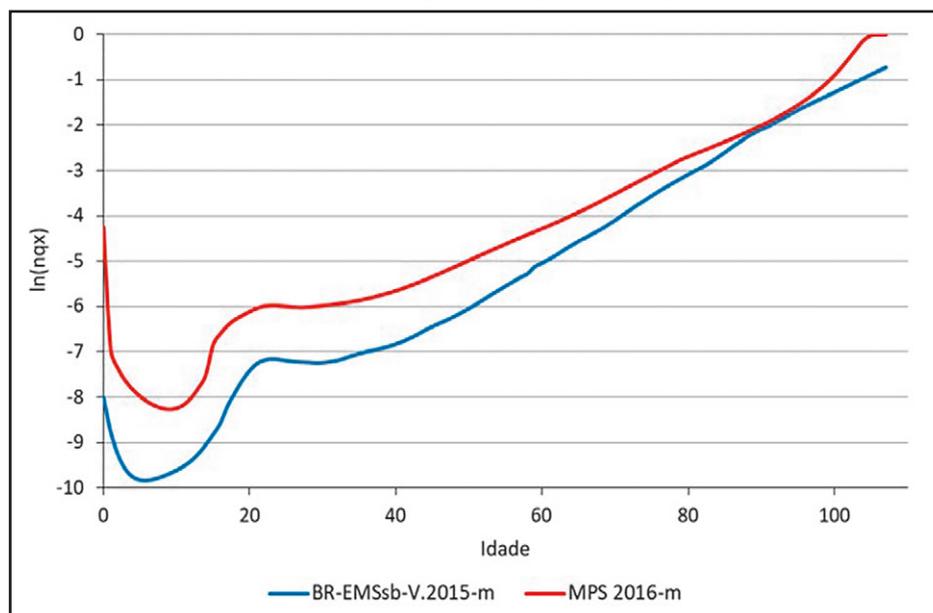
Vale ressaltar que, principalmente no caso de seguros privados, a população segurada tem características muito específicas que geralmente resultam numa esperança de vida consideravelmente maior da que caracteriza a população geral. Esta diferença é tão grande que até 2010 as companhias de seguro no Brasil preferiam usar tábuas de vida elaboradas para populações seguradas norte-americanas, por se aproximarem mais do padrão de mortalidade dos seus clientes brasileiros, do que as tábuas de vida calculadas pelo IBGE para a totalidade da população. A escolha dessas tábuas se dava por testes de aderências, pelos quais se determina, dentre um conjunto de tábuas de vida, qual melhor se adere à população em questão, comparando as mortes observadas na população segurada com as esperadas por cada tábua. Para exemplificar esta diferença nos níveis de mortalidade, o Gráfico 9.10 apresenta o logaritmo das probabilidades de mortes pela tábua AT 2000 masculina, a Annuity Table, elaborada a partir de contratos de

anuidades de companhias de seguros nos EUA. Como pode-se observar pelo gráfico, para quase todas as idades, em especial as mais avançadas, a AT 2000 apresenta mortalidade menor que a tábua IBGE 2016, apresentando, portanto, população mais longeva, razão pela qual ela muitas vezes foi utilizada no mercado segurador.

Wilbert, Lima e Gomes (2013) analisaram o impacto da utilização de diferentes tábuas de mortalidade nas estimativas do pagamento de benefícios no Regime Geral de Previdência Social (RGPS), usando duas tábuas distintas: a do IBGE, usada pelo Ministério da Previdência Social, e uma elaborada por Souza (2009), com base na mortalidade da população idosa aposentada pelo RGPS, para os anos de 1998-2002. Com base apenas no número de benefícios concedidos em 2002, a diferença da expectativa de vida entre as tábuas representou um impacto negativo de cerca de 4,5% do deficit previdenciário apurado pelo RGPS nesse mesmo ano, evidenciando a necessidade de se calcular tábuas de mortalidade específicas para a massa de segurados do RGPS. Outros esforços para construir tábuas de vida para populações previdenciárias específicas incluem o trabalho de Beltrão e Sugahara (2017), que construiu tábuas de vida para funcionários públicos civis do Executivo por sexo e escolaridade entre 1993 e 2014, atendendo a uma demanda da Fundação de Previdência Complementar do Servidor Público Federal, criada em 2013.

Como se trata de operações comerciais, as bases de dados sobre mortes de assegurados por esquemas privados costumam ser de melhor qualidade do que os registros de mortalidade da população em geral. Portanto, algumas vezes é possível construir tábuas de vida com os dados próprios dessas populações. Um exemplo é a família de tábuas BR-EMS, da Experiência do Mercado Segurador Brasileiro, adotada pela Federação Nacional de Previdência Privada de Vida (FenaPrevi), e agora parâmetro para o mercado segurador brasileiro. As tábuas BR-EMSsb 2015, por exemplo, foram construídas com dados de sobrevivência de 2007 a 2012 de mais de 39 milhões de indivíduos, aproximadamente 80% da população brasileira que possui seguros de vida privados com cobertura de sobrevivência no Brasil, ou seja, que dão direito a benefícios em vida (aposentadoria). Para ilustrar, o Gráfico 9.10 apresenta a tábua BR-EMSsb 2015 masculina e a compara às do IBGE 2016 e AT 2000, também masculinas. É notável que as probabilidades de morte em cada idade da população que tem seguro privado, dada pelas tábuas AT 2000 e BR-EMSsb 2015, são consideravelmente mais baixas do que as probabilidades da população masculina geral, dada pela tábua IBGE 2016. Contudo, a tábua BR-EMSsb 2015 apresenta estrutura de mortalidade mais similar à mortalidade brasileira como um todo, se diferenciando da AT 2000 principalmente nos primeiros grupos etários, o que justifica a elaboração de uma tábua de mortalidade específica para o mercado segurador brasileiro.

Gráfico 9.10: A função  $\ln(nqx)$  para a tábua BR-EMSsb-V.2015-m, comparada com a tábua de vida masculina do Brasil (MPS 2016-m) produzida pelo IBGE para 2016



Fonte: Oliveira et al. (2016): Appendix Table 1; IBGE (2016).

Embora as tábuas BR-EMSsb-V.2015-m e IBGE 2016 apresentem níveis de mortalidade muito diferentes, elas se aproximam nas idades entre 85 e 95 anos. Isso se dá pela diferença de metodologia adotada na extrapolação das duas tábuas. Para a tábua IBGE 2016 itera-se para encontrar um Fator de Ajuste  $FA$  na fórmula  $\ell_{x+2} = \ell_{x+1} \ell_{x+1} / (\ell_x + FA)$ , tal que a esperança de vida estimada pela tábua de vida completa seja igual à estimada pelo IBGE na tábua de vida com grupos etários abreviados, mantendo a coerência dos resultados (Castro, 2018). Já para a tábua BR-EMSsb-V.2015-m, para os grupos etários mais avançados, as funções  $q_x$  foram obtidas por extrapolações exponenciais dos  $q_x$  obtidos nos grupos intermediários (Oliveira et al., 2016). Essas metodologias procuram estimar, da melhor forma possível, a mortalidade nos últimos grupos etários retirando os efeitos de aleatoriedade pela pequena quantidade de dados. Contudo, ainda pouco se sabe sobre o comportamento da função de mortalidade nas idades mais elevadas, pois é recente a queda da mortalidade a ponto de permitir que uma grande quantidade de indivíduos chegasse até lá.

À parte dessa discussão, é muito comum, ainda, para os cálculos atuariais, a utilização da metodologia de tábuas de vida para medir outras formas de saída da população que não a morte. Em planos previdenciários, por exemplo, são utilizadas tábuas de saída da população de contribuintes por invalidez, aposentadoria ou demissão. Como o risco de sair por todas essas causas atuam sobre os mesmos indivíduos num mesmo tempo, utiliza-se a metodologia de Tabelas de Múltiplos Decrementos para considerar a forma de interação entre eles, conforme descrito no Capítulo 19. Em planos previdenciários também é comum o uso de tábuas de mortalidade diferentes para grupos populacionais diferentes. Aos inválidos, por exemplo, que em geral apresentam alguma doença que acelera sua mortalidade em relação aos não inválidos, geralmente é aplicada uma tábua de mortalidade diferente, construída exclusivamente para esta população (Pitacco, 2012; Ribeiro et al., 2007; Ribeiro, Reis e Barbosa, 2010). Beltrão et al. (2004) apresentam uma visão geral das diferentes tábuas de vida atuariais usadas no mercado brasileiro de seguros.

## 9.7 APLICAÇÕES DE FUNÇÕES DA TÁBUA DE VIDA

Nesse sentido, é importante ressaltar que a diminuição da mortalidade pode afetar a esperança de vida de formas muito diferentes, a depender da idade em que a queda acontece. Como apresentado na seção 2.2, no início da queda da mortalidade é comum observar grande diminuição da mortalidade infantil, pois em geral as causas de mortalidade em populações em que a mortalidade é muito alta estão relacionadas à falta de saneamento básico e doenças infectocontagiosas. A melhoria dessas condições afeta principalmente as crianças, resultando em grande redução da mortalidade infantil. Uma vez que essas crianças superam o risco de mortalidade nos primeiros grupos etários, é provável que venham a morrer apenas nas idades adultas ou idosas, de forma que passam a contribuir com uma grande quantidade de anos-pessoa para a tábua de vida.

A redução da mortalidade que ocorre nas idades adultas ou idosas causa um aumento do número de anos vividos que é menor do que o efeito de uma redução equivalente na mortalidade das crianças. Uma consequência disso é que quando a redução da mortalidade se dá nas primeiras idades, seu efeito na esperança de vida ao nascer é muito maior que quando se dá nas idades mais avançadas. Em contrapartida, a queda na mortalidade infantil pode não afetar significativamente a esperança de vida nas idades mais avançadas, pois não representa mudança na intensidade com que as pessoas morrem naquelas idades. De forma similar, a queda na mortalidade nas idades mais avançadas em decorrência da queda na mortalidade por doenças degenerativas e induzidas pela atuação humana não tem grande efeito na esperança de vida ao nascer, mas pode afetar consideravelmente as esperanças de vida mais avançadas.

Essa reflexão tem algumas aplicações práticas interessantes, entre elas, a discussão sobre reformas previdenciárias. Se um benefício previdenciário é recebido vitaliciamente a partir dos 60 anos, por exemplo, o tempo esperado de recebimento deste benefício corresponde à esperança de vida aos 60 anos, função que se altera no tempo de forma diferente da esperança de vida ao nascer, dependendo da forma como se dão as diversas mudanças na função de mortalidade. Portanto, para analisar a variação no tempo de recebimento deste benefício entre dois períodos distintos, deve-se analisar a diferença na esperança de vida a partir dos 60 anos nesses dois períodos, e não a diferença na esperança de vida ao nascer entre as duas datas.

Para a análise da população beneficiária em relação à população contribuinte as funções da tábua de vida também são úteis. Assumindo que as pessoas contribuem para a previdência entre as idades 30 e 60 anos, pode-se dizer que, se a população experimenta a mortalidade dada por certa tábua de vida, a diferença  $(T_{30}-T_{60})$  representaria o tempo total vivido pela coorte em idade de realizar contribuição, enquanto a função  $T_{60}$  representaria o tempo recebendo benefícios. Assim,  $T_{60}/(T_{30}-T_{60})$  representa a relação entre o tempo que se passa recebendo benefício e o tempo que se passa contribuindo para a previdência, em média, numa população, assumindo que todos contribuem e recebem benefícios. Pelos dados do Gráfico 9.7, por exemplo, a razão  $T_{60}/(T_{30}-T_{60})$  é igual a 0,7 para mulheres e 0,6 para homens, indicando que mulheres passam mais tempo recebendo benefícios em relação a seu tempo de contribuição que homens, assumindo que homens e mulheres têm a mesma condição de ingressar e permanecer no mercado de trabalho, e que todos contribuem durante toda sua vida ativa, e recebem benefícios a partir dos 60 anos.