

Geração de um modelo de tábuas de vida representativo da mortalidade do Estado de São Paulo*

Neir Antunes Paes**

É tentador especular se um sistema de quatro parâmetros para modelar o padrão de mortalidade de uma população pode ser aplicado consistentemente em regiões do país com dados de boa qualidade e, assim, lançar algumas luzes sobre os padrões de mortalidade de outras regiões com dados de qualidade duvidosa.

O uso de métodos indiretos para estimar indicadores da mortalidade apóia-se em suposições sobre o padrão aproximado de mortalidade específica por idade da região em estudo. Para este propósito, modelos de tábuas de vida têm sido identificados, os quais fornecem típicos padrões de mortalidade. A maior parte desses modelos foi baseada exclusivamente em populações de origem européia, o que pode torná-los inapropriados para representar a experiência de mortalidade de regiões em desenvolvimento. Uma vez que a estrutura de mortalidade da população brasileira ainda é desconhecida, devido à qualidade deficiente das suas estatísticas de óbitos, a aplicação de um modelo padrão adequado para

o Brasil tem sido freqüentemente tratada como uma questão de adivinhação. Não existe consenso sobre qual modelo seja o mais apropriado para as regiões brasileiras, o que pode conduzir a duvidosas estimativas de mortalidade baseadas em métodos indiretos.

No Brasil, este problema foi discutido pela primeira vez pelo IBGE (1974) e por Frias (1974). Observou-se, então, que o padrão da mortalidade brasileira podia diferir sistematicamente dos conjuntos de tábuas de vida tradicionais, como o de Coale e Demeny (1966) e o das Nações Unidas (1982).

A necessidade de construir modelos de tábuas de vida com a real experiência de mortalidade brasileira levou

* Este artigo consiste em um capítulo da minha tese de doutorado, *Model life table representation for Brazilian mortality*, defendida na Universidade de Londres em 1993.

** Professor adjunto do Departamento de Estatística da Universidade Federal da Paraíba (UFPB), Campus I, João Pessoa.

Frias e Rodrigues (1981) a desenvolverem um sistema de tábuas de vida com um parâmetro por sexo, designado de modelo Brasil, o qual se propõe a representar os padrões de mortalidade das regiões brasileiras. Este estudo pioneiro proporcionou um grande avanço inicial nesta área mas, embora útil (na ausência de outras propostas), o modelo parece não ser adequado para descrever as diferentes experiências de mortalidade de subpopulações nacionais. As distintas regiões do Brasil estão em diferentes níveis de desenvolvimento e processos de mudanças, os quais podem determinar padrões de mortalidade individuais e diferenciados. Do mesmo modo que os modelos clássicos de tábuas de vida – o de Coale e Demeny (1966) e o das Nações Unidas (1982) – e o sistema logito de dois parâmetros proposto por Brass (1975), o modelo Brasil também parece ser insuficientemente flexível para modelar adequadamente os padrões de mortalidade observados no país.

Dois sistemas de quatro parâmetros para modelar o padrão de mortalidade observado em uma população, Zaba (1979) e Ewbank *et al.* (1983), foram desenvolvidos para situações em que o sistema logito de dois parâmetros não é suficiente. Talvez por requererem dados de alta qualidade, estes sistemas ainda não foram usados no Brasil. É tentador especular se eles podem ser aplicados consistentemente em regiões do país com dados de boa qualidade e, assim, lançar algumas luzes sobre os padrões de mortalidade de outras regiões com dados de qualidade duvidosa.

Neste artigo são feitas tentativas para encontrar um modelo padrão adequado para representar a experiência de mortalidade do Estado de São Paulo. Um novo modelo é proposto usando o siste-

ma adaptativo de quatro parâmetros de Ewbank *et al.* (1983).

Utilização de sistemas de tábuas de vida

Possivelmente, a técnica mais simples para ajustar um conjunto de probabilidades de sobrevivências – valores de $\ell(x)$ – é fornecida pelo sistema logito (1). Para o seu uso, exige-se a seleção de um modelo padrão de tábuas de vida que seja o mais próximo possível das tábuas de vida observadas.

Vários conjuntos de tábuas de vida têm sido propostos. Contudo, o modelo geral de Brass é sempre uma possibilidade a ser considerada, uma vez que ele revela um padrão muito similar ao modelo Oeste do conjunto de tábuas de vida de Coale e Demeny, usado freqüentemente para representar a mortalidade de países em desenvolvimento. O modelo geral de Brass representa um padrão geral de mortalidade, e os dados que o geraram foram provenientes de países com alta mortalidade infantil relativamente à mortalidade jovem e adulta.

Tábuas de vida do Estado de São Paulo de 1940 a 1970 (em períodos decenais) foram incluídas no conjunto de tábuas de vida usadas na construção do modelo Brasil, o que o torna um candidato natural a ser considerado no estudo dos modelos relacionais.

Usou-se o sistema logito de dois parâmetros e fez-se uma regressão das transformações do logito, $Y(x)$, de cada $\ell(x)$ observada com os logitos do modelo geral de Brass. O mesmo foi feito com relação ao modelo Brasil, cujas probabilidades de sobrevivências $\ell(x)$ foram encontradas adotando-se o nível 18

(1) O logito de um conjunto de valores de sobrevivência é definido por: $Y(x) = \text{logito}[\ell(x)] = 0,5 \cdot \log_e \left[\frac{1 - \ell(x)}{\ell(x)} \right]$.

como modelo central para ambos os sexos.

Quando os dados são fidedignos, existem muitas possibilidades de identificação das características do padrão de mortalidade por idade. Os ajustes dos modelos de quatro parâmetros para dados observados pressupõem o conhecimento do padrão de mortalidade, especialmente para as idades extremas. Afortunadamente, a qualidade dos registros de óbitos do Estado de São Paulo permite a utilização de modelos sofisticados. Desta forma, os dois sistemas de quatro parâmetros – o de Zaba e o de Ewbank *et al.* – são também usados neste trabalho.

A qualidade dos dados disponíveis

Os dados de população por idade e sexo do Estado de São Paulo foram extraídos dos censos demográficos, e os registros de óbitos, de Ferreira (1980). O conjunto de dados de população e óbitos estão disponíveis a partir de 1940 até 1980, para cada ano decenal, permitindo a construção de dez tábuas de vida para São Paulo por sexo.

A série de 40 anos de experiência da mortalidade de São Paulo reflete o período mais importante da transição da mortalidade no Estado. A Tabela 1 mostra sua evolução em termos da esperan-

ça de vida ao nascer e da taxa de mortalidade infantil. A vida média observada aumentou rapidamente de 44,5 anos para 62,5 anos, para o sexo masculino, e de 46,8 para 68,7 anos para o sexo feminino. Este aumento representa um ganho de 40,2% (média anual de 0,45) para o sexo masculino e de 46,6% (média anual de 0,55) para o sexo feminino. A mortalidade infantil para o sexo masculino declinou de 161,1 para cada mil nascidos vivos em 1940 para 61,7 em 1980, uma redução de cerca de 62%. Para o sexo feminino esta redução foi maior, 67%, expressando o declínio da mortalidade infantil de 143,5 para 47,8 por mil.

A análise da qualidade da distribuição da população por idade e sexo em São Paulo no período de 1940 a 1980 revelou que os erros não são suficientes para justificar correções na estrutura por idade.

Por sua vez, a qualidade dos registros dos óbitos em relação à cobertura dos censos foi extensivamente analisada por Paes (1993), que encontrou um sub-registro inferior a 5% para ambos os sexos nas idades de cinco anos e mais no período 1940-80. Com base nestes achados, assumiu-se que os registros de óbitos são confiáveis e decidiu-se não fazer nenhuma correção.

Evidências mostrando que o registro de óbitos para as idades abaixo dos

Tabela 1
Esperança de Vida ao Nascer (e^0) e Taxa de Mortalidade Infantil (TMI) geradas dos modelos adaptativos para São Paulo 1940-1980

Ano	Masculino						Feminino					
	e^0			TMI			e^0			TMI		
	OBS	EWB-4	SP-2	OBS	EWB-4	SP-2	OBS	EWB-4	SP-2	OBS	EWB-4	SP-2
1940	44,5	44,5	44,5	161,1	176,5	176,0	46,8	46,8	46,8	143,5	153,6	151,0
1950	52,9	52,9	52,9	116,3	122,9	124,9	55,8	55,8	55,8	101,7	110,3	111,1
1960	59,0	59,0	59,0	78,2	78,5	80,4	63,2	63,2	63,2	65,7	67,9	72,4
1970	58,5	58,5	58,5	97,2	88,6	87,0	65,0	65,0	65,0	70,2	71,1	68,3
1980	62,5	62,5	62,5	61,7	57,5	56,8	68,7	68,7	68,7	47,8	46,7	46,5

cinco anos é quase completo podem ser encontradas em Puffer e Serrano (1973), para o ano de 1970, e comentários neste sentido foram feitos, por exemplo, em Altman e Ferreira (1979) pelo menos desde 1940. Decidiu-se, contudo, usar os ajustamentos realizados por Ferreira (1980) para o período 1940-70 para as idades abaixo de dois anos. Similarmente, adotou-se o mesmo procedimento para 1980.

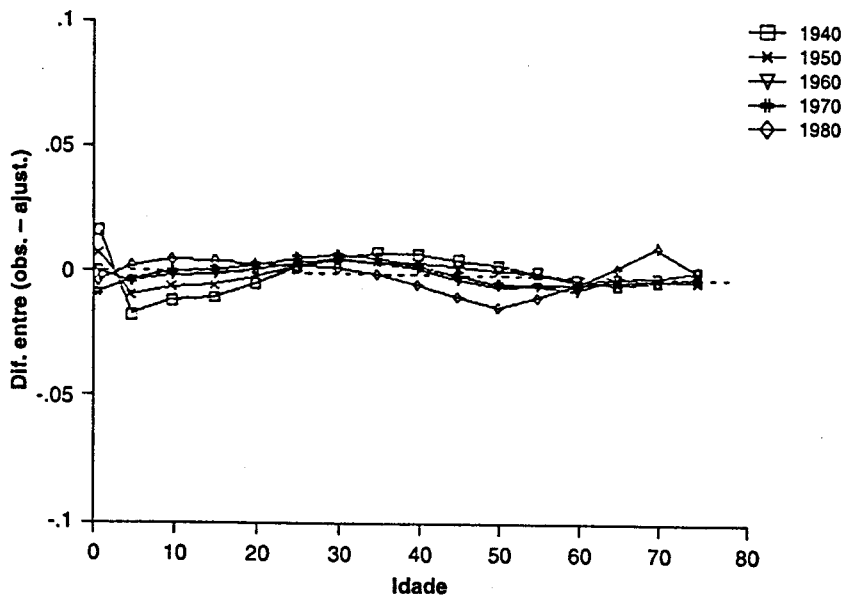
As taxas específicas de morte foram calculadas e ajustadas, sempre que necessário (2), usando-se a média dos três anos de registros de óbitos em torno dos anos censitários em grupos quinquenais (exceto para os grupos 0-1 e 1-4

anos). Os dados de população correspondentes por idade e sexo foram ajustados para o meio de período usando a taxa de crescimento intercensitário.

"Bondade" dos ajustamentos

Os quatro modelos citados acima foram aplicados à série de dados para São Paulo e os valores de $l(x)$ foram ajustados até a idade de 75 anos. Os Gráficos de 1 a 8 servem como guia para indicar a variabilidade dos valores de $l(x)$ ajustados, os quais são comparados com os valores de $l(x)$ observados.

Gráfico 1
Diferenças entre $l(x)$ Obs./Ewbank
1940-80 - Masculino



(2) As taxas estão assinaladas em termos dos logaritmos das probabilidades de morte acima da idade de dois anos e as discrepâncias foram ajustadas por inspeção visual, principalmente para as idades mais avançadas.

Gráfico 2
Diferenças entre $\ell(x)$ Obs./Zaba
1940-80 - Masculino

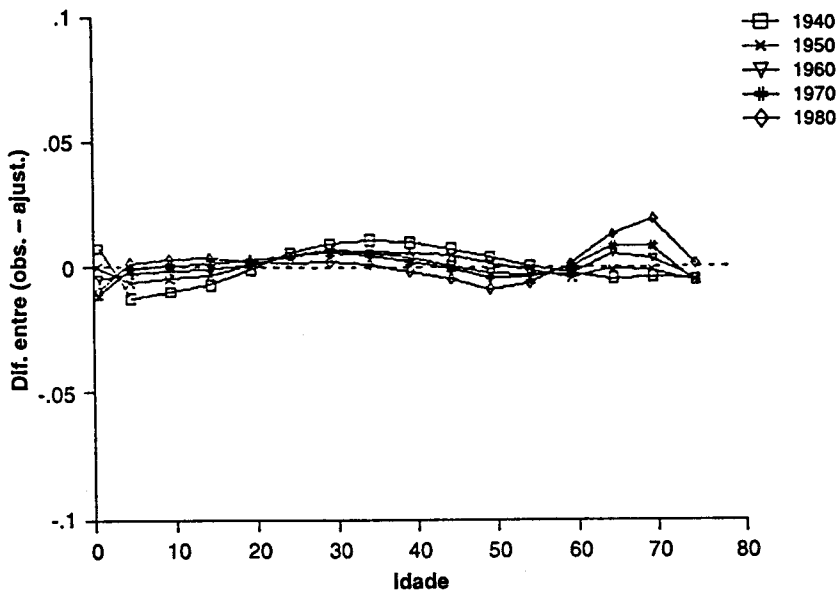


Gráfico 3
Diferenças entre $\ell(x)$ Obs./Brass
1940-80 - Masculino

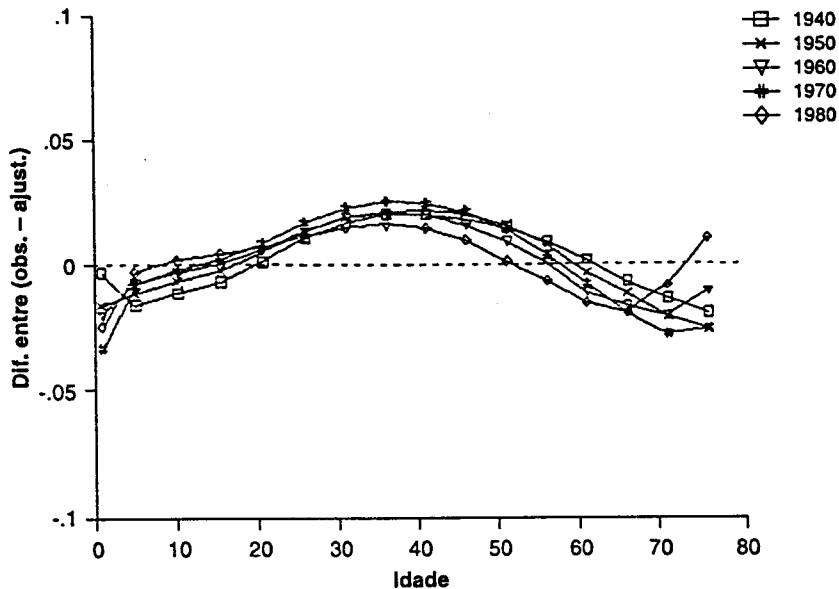


Gráfico 4
Diferenças entre $\ell(x)$ Obs/M-Brasil
1940-80 - Masculino

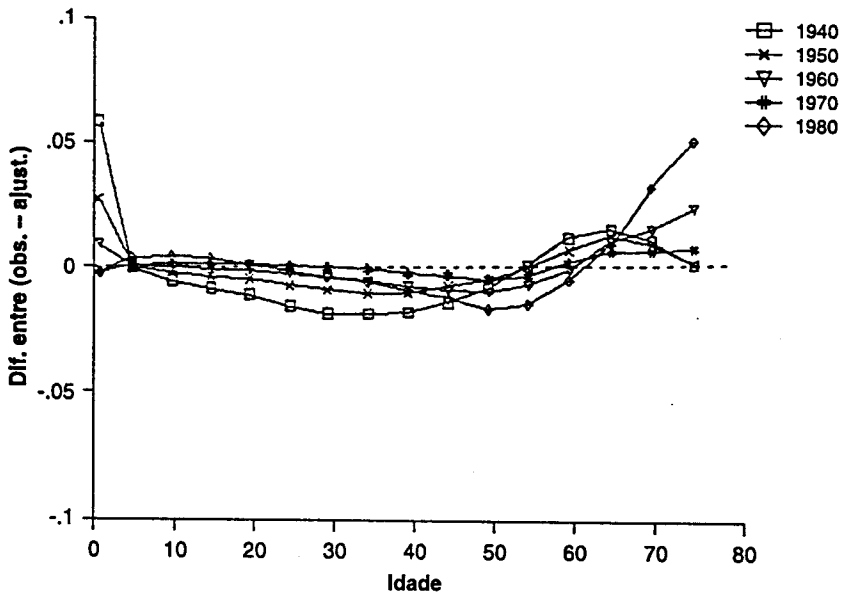


Gráfico 5
Diferenças entre $\ell(x)$ Obs/Ewbank
1940-80 - Feminino

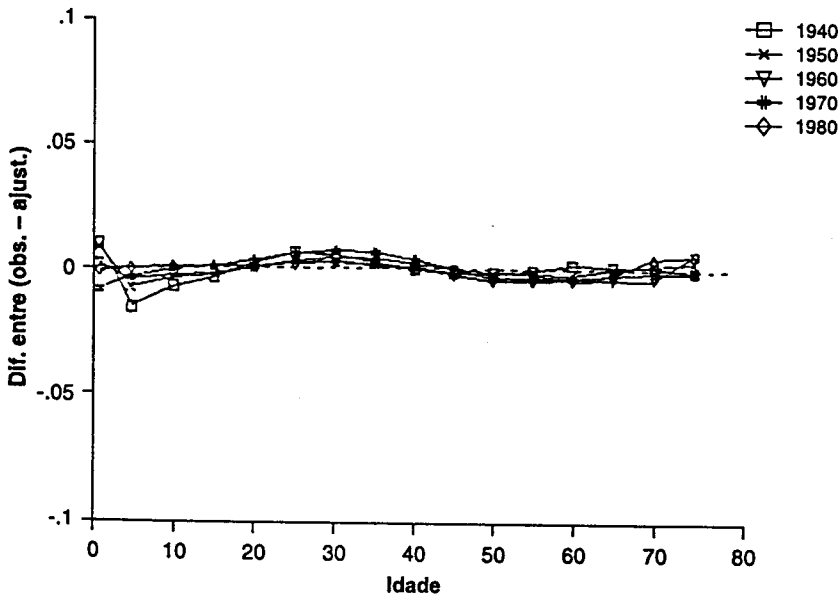


Gráfico 6
Diferenças entre $\ell(x)$ Obs./Zaba
1940-80 - Feminino

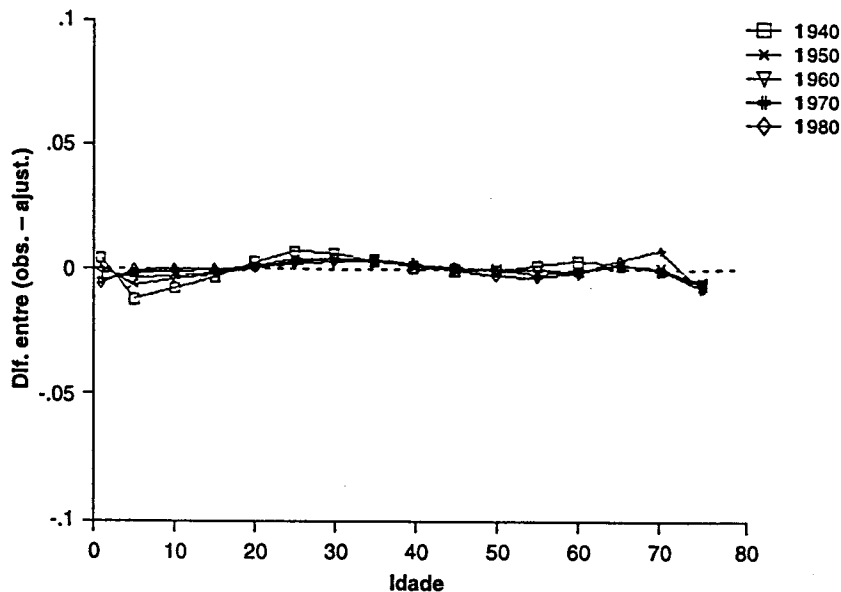


Gráfico 7
Diferenças entre $\ell(x)$ Obs./Brass
1940-80 - Feminino

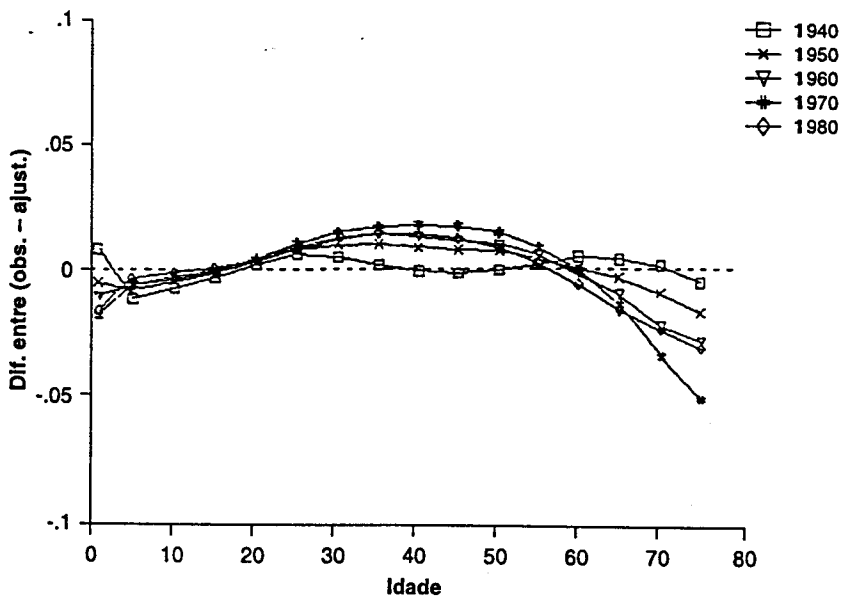
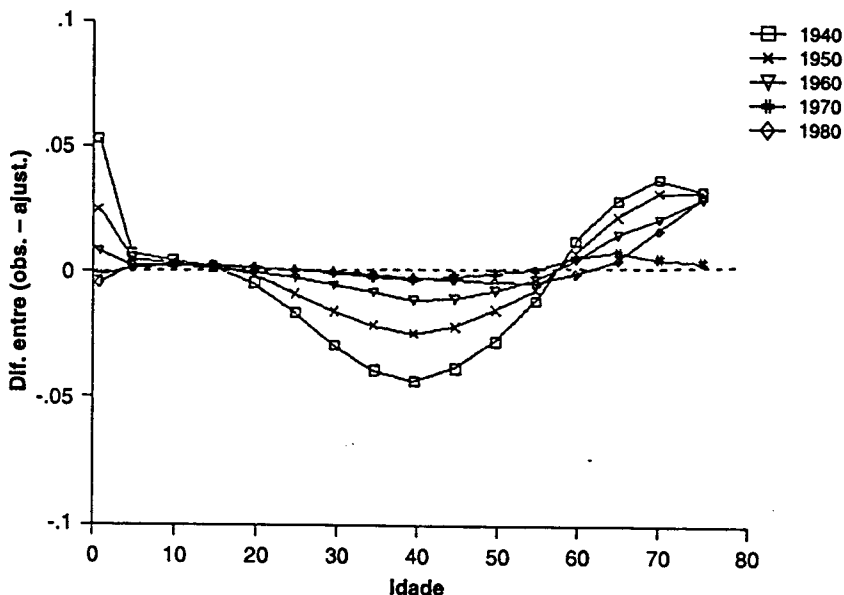


Gráfico 8
Diferenças entre $\ell(x)$ Obs/M-Brasil
1940-80 - Feminino



Os resultados destes ajustamentos indicam a capacidade de cada modelo para ajustar os valores de $\ell(x)$. Como alguns modelos podem capturar a complexidade da curva melhor que outros, é necessário usar alguma função objetiva como uma medida-resumo da "bondade do ajustamento". Utiliza-se, com este fim, o EQM (Erro Quadrático Médio) para os valores de $\ell(x)$ por idade da tábua de vida e para os seus logitos. Para $\ell(x)$ a estatística adotada é:

$$EQM[\ell] = \sqrt{\frac{\sum_{x=1}^{75} [\ell(x) - \ell_s(x)]^2}{n}}$$

onde n é o número de pontos correspondente aos grupos de idades usados na

comparação; $\ell(x)$ é a probabilidade de sobrevivência observada desde o nascimento; e $\ell_s(x)$ é o valor de $\ell(x)$ estimado proveniente do modelo. A definição para $Y(x)$ é similar.

Quando o modelo de Brass é usado, os gráficos das diferenças, para ambos os sexos (Gráficos 3 e 7), mostram que (exceto para o sexo feminino em 1940) grandes diferenças ocorrem nas idades 1, 25-50 e 65 e mais, sugerindo que o modelo ajusta precariamente as tábuas de vida observadas para São Paulo.

Quando o modelo Brasil é aplicado (Gráficos 4 e 8), os ajustes para 1970 são muito próximos para ambos os sexos, mas o tamanho das diferenças demonstra que o modelo ajusta as tábuas de vida precariamente para os outros anos. Tais diferenças são maiores para o sexo feminino e para as idades extremas e inter-

Tabela 2
Comparação dos ajustes de cinco modelos de tábuas de vida para São Paulo
1940-1980

Ano	EQM de $Y(x)$					EQM de $\ell(x)$				
	BRASS	M-BRA	ZABA	EWB-4	SP-2	BRASS	M-BRA	ZABA	EWB-4	SP-2
Masculino										
1940	.034500	.054200	.018200	.022900	.035500	.013400	.018600	.007200	.008200	.014900
1950	.045200	.034400	.012900	.015700	.024700	.016000	.009600	.004400	.004400	.008500
1960	.058800	.026600	.015500	.012200	.020500	.014400	.008900	.003800	.003300	.005000
1970	.077100	.008800	.021700	.018700	.038800	.019400	.003400	.005200	.004200	.009300
1980	.078500	.043400	.029900	.023100	.029700	.012600	.016400	.007100	.006100	.005300
Todos	.058800	.033500	.019600	.018500	.029800	.015200	.011500	.005500	.005200	.008600
Feminino										
1940	.014800	.076100	.014000	.017300	.017300	.005100	.028800	.005100	.005700	.006400
1950	.022500	.052900	.010400	.015300	.014900	.007900	.018200	.003100	.003800	.003700
1960	.043700	.033500	.011900	.014300	.016600	.012400	.011000	.002800	.003500	.004100
1970	.067900	.010000	.014700	.009700	.014500	.019600	.002700	.003100	.002800	.002800
1980	.074200	.025800	.020900	.012700	.015100	.013800	.008600	.003600	.003200	.003900
Todos	.044600	.039700	.014400	.013900	.015680	.011800	.013900	.003500	.003700	.004140

mediárias. Estes ajustes sugerem que o modelo Brasil não captura toda a complexidade do padrão de mortalidade de São Paulo para os diferentes anos. O fato dele ajustar bem os valores de $\ell(x)$ para 1970 pode ser devido à forte representação de tábuas de vida para este ano, somada ao peso da série histórica de tábuas de vida do Estado de São Paulo utilizada como base de dados na construção do modelo. Assim sendo, reforça-se a necessidade de recorrer a modelos alternativos para ajustar as tábuas de vida para o Estado de São Paulo.

Os modelos de Zaba (Gráficos 2 e 6) e de Ewbank *et al.* (Gráficos 1 e 5) mostram padrões de diferenças similares para ambos os sexos. Estas diferenças são relativamente pequenas para todas as idades, particularmente para o sexo feminino. Uma vez que ambos os modelos produzem ajustamentos muito próximos, é difícil identificar qual deles fornece o melhor ajustamento apenas

observando os gráficos e os desvios dos valores.

A Tabela 2 sintetiza os desvios em termos de EQM para cada um dos quatro modelos utilizados para ajustar as tábuas de vida observadas. Os ajustes obtidos pelos dois sistemas de quatro parâmetros são bons se comparados com aqueles produzidos pelo modelo de Brass e o modelo Brasil. O EQM para ambos os índices, $\ell(x)$ e $Y(x)$, demonstra claramente que a boa *performance* dos dois sistemas vale para todos os anos investigados (3).

Quando os valores médios do EQM, em termos de $Y(x)$, são comparados para todos os anos nos dois sistemas de quatro parâmetros, eles apresentam-se menores para o sistema de Ewbank para ambos os sexos. O EQM em termos de $\ell(x)$ para ambos os sexos virtualmente não revela nenhuma diferença em ambos os sistemas.

Uma vantagem do modelo adaptativo de Ewbank é que o procedimento de

(3) Exceto o sistema de Ewbank para o sexo feminino em 1940, quando comparado com o de Brass. Isto ocorreu porque, na aplicação do sistema de Ewbank, foi utilizado um novo modelo padrão desenvolvido pelos próprios autores. Usando-se o modelo geral de Brass no sistema de quatro parâmetros, os valores do EQM seriam 0,0129 para $Y(x)$ e 0,0045 para $\ell(x)$.

ajuste depende de dados para as primeiras e últimas idades, ao passo que o sistema de Zaba exige o conhecimento de valores de sobrevivência para idades individuais. Como o sistema de Ewbank é mais flexível que o de Zaba e ajusta melhor, o primeiro será adotado, a partir de agora, como o modelo mais adequado para ajustar as tábuas de vida do Estado de São Paulo (ver Apêndice).

Os dados da Tabela 1 – que mostra as esperanças de vida ao nascer e as taxas de mortalidade infantil obtidas das tábuas de vida observadas e ajustadas pelo sistema de Ewbank para cada ano em ambos os sexos –, quando comparados com os valores observados, mostram que as expectativas de vida são exatamente as mesmas e as taxas de mortalidade, muito próximas. Com exceção do ano de 1940, as diferenças entre as duas fontes de estimativa foram abaixo de dez crianças para cada mil nascidas vivas.

Em busca de um modelo representativo melhor

O uso do modelo adaptativo de Ewbank pressupõe um completo conhecimento do padrão de mortalidade, especialmente das idades extremas. Na impossibilidade de satisfazer a esta exigência, não é possível usar este procedimento. Diante destas restrições, especula-se que seria mais prático o uso de um modelo de tábuas de vida que exigisse menos parâmetros, mas que produzisse ajustes os mais próximos possíveis daqueles obtidos quando os quatro parâmetros são usados.

Investigações preliminares tendo em vista tais reparametrizações foram

feitas por Ewbank *et al.* (1983) usando tábuas de vida da Suécia. Registra-se, ainda, uma tentativa feita com tábuas de vida de países da América Latina, que não teve sucesso. A sugestão do autor foi a de reduzir o modelo de quatro parâmetros a um modelo com dois ou três parâmetros, tentando encontrar relacionamentos significativos entre eles.

Inter-relacionamentos entre os quatro parâmetros do modelo de Ewbank

A Tabela 3 mostra os valores estimados de cada parâmetro, para ambos os sexos, obtidos pela aplicação do modelo de Ewbank para o período 1940-80. Os valores de alfa (4) e beta se referem às estimativas obtidas na primeira iteração. As estimativas revelam que a maioria dos valores dos parâmetros muda monotonicamente. Estas tendências são quebradas somente para o ano de 1970 para o sexo masculino, quando o nível da mortalidade de São Paulo permaneceu quase o mesmo de 1960.

Fazendo-se a regressão dos valores de α , obtidos na primeira iteração, com os valores de α da última iteração para ambos os sexos e, por sua vez, a regressão dos primeiros valores de β com os últimos por sexo, produziram-se os seguintes resultados em termos de R^2 (Coeficiente de Determinação). A regressão dos valores de α produziu um $R^2 = 98,3\%$ e $99,4\%$ para os sexos masculino e feminino, respectivamente. Estas correlações quase perfeitas entre os valores dos alfas e betas sugerem que não existem diferenças significativas entre a primeira iteração e as subsequentes. A razão para usar a primeira iteração é que

(4) O sistema de Ewbank fornece valores positivos de alfa. Isto ocorre porque os autores usam logito $[r(x)] = 0,5 \log_e [r(x)/(1-r(x))]$ para calcular $Y(x)$. Aqui os valores de alfa são apresentados na forma convencional, ou seja, com sinal negativo.

Tabela 3
Estimativas dos parâmetros usando o modelo Ewbank para São Paulo
1940-1980

Ano	Masculino				Feminino			
	alfa	beta	kappa	lambda	alfa	beta	kappa	lambda
1940	-0,051	0,987	-0,170	-0,064	-0,107	0,844	0,066	-0,066
1950	-0,314	1,095	-0,400	0,048	-0,399	0,918	-0,098	-0,016
1960	-0,530	1,290	-0,570	0,180	-0,685	1,000	-0,370	-0,007
1970	-0,520	1,270	-0,663	0,166	-0,780	1,060	-0,750	0,024
1980	-0,683	1,560	-0,803	0,367	-0,970	1,240	-0,890	0,180

se formos usar um modelo adaptativo de dois parâmetros para ajustar dados menos confiáveis – do qual não se tem todos os valores de $\ell(x)$ para um ajustamento através do uso dos quatro parâmetros – não será possível encontrar valores iterativos de alfa e beta.

A Tabela 4 revela que as correlações entre os parâmetros em termos de R^2 são muito altas; com exceção dos relacionamentos entre lambda e alfa para o sexo feminino, todos os outros valores são pelo menos da ordem de 92%. Para ambos os sexos, as regressões de κ com α e λ com β são as mais elevadas, especialmente para o sexo masculino. Os relacionamentos entre estes parâmetros, cujos R^2 foram os mais altos, podem ser vistos nos Gráficos de 9 a 12. Eles mostram claramente vários padrões.

Tabela 4
Coefficiente de Correlação (R^2) segundo o
relacionamento entre os parâmetros por
sexo

Relacionamento entre os Parâmetros	R^2	
	Masculino	Feminino
κ e α	0,97	0,92
κ e β	0,93	0,93
λ e α	0,93	0,72
λ e β	0,99	0,94

De acordo com estes resultados, é possível dizer que, historicamente, o modelo de tábuas de vida ajustado para São

Paulo de 1940 a 1980, para ambos os sexos, revela um forte relacionamento entre os parâmetros que governam a mortalidade nas idades idosas (λ) e o padrão da mortalidade (β). Também mostra um forte relacionamento entre o parâ-

Gráfico 9
Relacionamento entre alfa/kappa
São Paulo, 1940-80 – Masculino

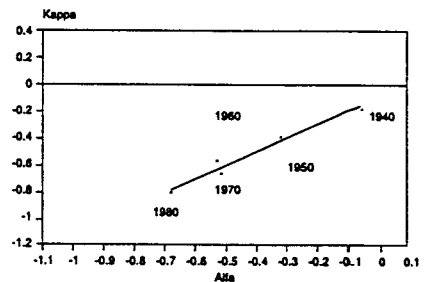


Gráfico 10
Relacionamento entre alfa/kappa
São Paulo, 1940-80 – Feminino

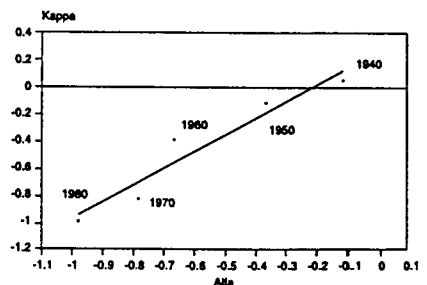


Gráfico 11
Relacionamento entre beta/lambda
São Paulo, 1940-80 - Masculino

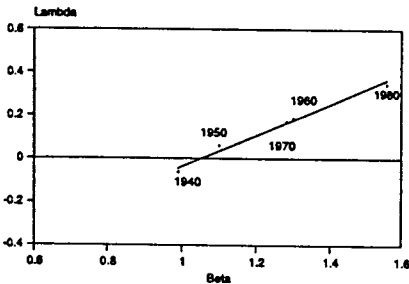
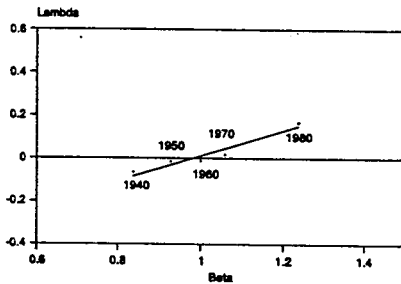


Gráfico 12
Relacionamento entre beta/lambda
São Paulo, 1940-80 - Feminino



metro que governa a mortalidade nas idades jovens (κ) e o nível da mortalidade (α). Estes fortes relacionamentos sugerem que seria possível reduzir o modelo de quatro parâmetros em um outro com apenas dois parâmetros. A regressão entre os pares de parâmetros cujas correlações foram as mais altas produziu os seguintes pares de equações lineares:

Para o sexo masculino,
 $\kappa = 0,1061 - 0,9892\alpha$ com $R^2 = 97\%$
 $\lambda = -0,7730 + 0,7356\beta$ com $R^2 = 99\%$

Para o sexo feminino,
 $\kappa = 0,2738 - 1,1598\alpha$ com $R^2 = 92\%$
 $\lambda = -0,5836 + 0,5992\beta$ com $R^2 = 94\%$

O relacionamento entre λ e β define a forma básica da curva nas idades ido-

sas. Uma vez que os quatro parâmetros resumem a forma da tábua de vida, eles podem ser usados para comparar os padrões de séries históricas de tábuas de vida.

Correlações cruzadas foram computadas para todos os pares de variáveis (alfa, beta, kappa e lambda) para todos os possíveis lags (Johnston, 1987). Uma ilustração é dada a seguir, comparando-se os pares de variáveis cujas correlações foram as mais altas (kappa contra alfa para o sexo masculino).

Lag (em anos)	Correl. cruzada (%)	-1	0	1
		+	+	+
0	97		∴	*
10	27		∴	*
20	-16		* ∴	
30	-28		* ∴	
40	-38		* ∴	

As correlações cruzadas entre as variáveis de lag zero foram de 97%, como mostra a Tabela 4. Para os outros lags, as mais altas correlações cruzadas entre os valores dos parâmetros (kappa e alfa) não excedem a -38% para o lag 40; isto é, 38% do relacionamento foi devido à correlação entre o valor de kappa em 1980 e o valor de alfa há 40 anos atrás, em 1940. No entanto, somente um ponto foi usado nesta correlação, de forma que este resultado não faz sentido.

Como esperado, as correlações cruzadas entre os outros pares de variáveis (kappa contra alfa para o sexo feminino e lambda contra beta para os sexos masculino e feminino) produziram correlações mais baixas para todos os lags. Note-se, contudo, que o reduzido número de pontos não permite que se façam generalizações. No entanto, o teste feito aqui sugere que é justificado o uso de um modelo de regressão linear entre os pares de variáveis, no qual se assume um

erro $\varepsilon_t \sim N(0, \sigma^2)$. Assim sendo, pode-se concluir por um relacionamento direto entre os parâmetros κ/α e λ/β , e não por um relacionamento devido a uma dependência conjunta no tempo.

Exemplo do procedimento de ajuste proposto

Na Tabela 5 ilustra-se a aplicação do sistema adaptativo de dois parâmetros São Paulo, chamado de SP-2, usando-se a tábua de vida para o sexo feminino em 1970.

O sistema adaptativo usa as razões de sobrevivência, $\ell(x)$, para pelo menos três idades adultas (Ewbank recomenda as idades intermediárias - 45, 50 e 55 anos). Os logitos desses valores (coluna 3) são usados para obter uma primeira estimativa dos parâmetros α e β fazendo a regressão em termos dos valores de $Y_s(x)$ (coluna 4). O padrão de mortalidade usado aqui consiste de uma variação do padrão geral de Brass, o qual foi elaborado por Ewbank *et al.* (1983). O segundo passo é estimar o valor de κ usando a equação dada acima. Do mesmo modo, obtém-se uma estimativa de λ . Ambas as estimativas encontram-se na coluna 6.

Calculando-se os valores de $T[\ell_s(x); \kappa, \lambda]$ (cuja equação é dada no Apêndice) para todos os grupos de idade (coluna 7), e fazendo a regressão deles com os valores observados $Y(x)$, obtêm-se estimativas revisadas para α e β , α_1 e β_1 . Repetindo este procedimento (5), novas estimativas podem ser obtidas para κ e λ (coluna 12). Duas iterações são suficientes para encontrar um ajustamento muito próximo dos dados originais. As

colunas 10, 11, 16 e 17 dão um diagnóstico da "bondade" dos ajustamentos em termos dos quadrados das diferenças.

Performance do modelo SP-2 para representar os dados

Usando as estimativas dos parâmetros baseadas nas equações de regressões dadas anteriormente, calculou-se o EQM para todas as cinco tábuas de vida por sexo para São Paulo quando o modelo adaptativo SP-2 foi aplicado. Os resultados, constantes da Tabela 2, mostram que o ajuste do modelo SP-2 (EQM = 0,298 e 0,0086 para o sexo masculino e 0,0155 e 0,0042 para o sexo feminino, respectivamente para $Y(x)$ e $\ell(x)$) é muito melhor que o do modelo de Brass, e melhor que o do modelo Brasil.

Comparado com o modelo de quatro parâmetros de Zaba para o sexo feminino, os resultados mostram uma pequena vantagem para o modelo Zaba (EQM = 0,0144 e 0,0035 para $Y(x)$ e $\ell(x)$, respectivamente). Os valores do EQM obtidos quando o modelo SP-2 foi aplicado são, na verdade, bastante próximos daqueles encontrados para o modelo adaptativo de quatro parâmetros de Ewbank para o sexo feminino (EQM = 0,0139 e 0,0037, respectivamente para $Y(x)$ e $\ell(x)$).

A estimativa da mortalidade infantil gerada a partir do uso deste novo modelo é mostrada na Tabela 1. Os resultados são similares àqueles estimados pelo modelo de quatro parâmetros. A maior diferença nos resultados entre os modelos é encontrada para o sexo feminino em 1960. Mesmo assim, esta diferença é menos que cinco óbitos por mil crianças nas-

(5) Em situações em que não existe disponibilidade de todos os grupos etários de $\ell(x)$, apenas uma interação deverá ser feita, uma vez que, para obter estimativas revisadas dos parâmetros, é necessário fazer a regressão dos valores de $T[\ell(x); \kappa, \lambda]$ com todos os valores $Y(x)$.

cidas vivas. O modelo de Ewbank, como esperado, aproximou-se mais das taxas observadas que o modelo SP-2. Entretanto, a diferença média absoluta no ajuste é menos que dois por mil. Em termos de esperança de vida, ambos os modelos produziram exatamente os mesmos valores observados. Este fato apenas confirma o caráter adaptativo dos modelos, ajustando-se muito bem aos dados observados para todos os anos, conforme já demonstrado. Do mesmo modo, o sistema SP-2 também produziu ajustes com a mesma aproximação.

Em termos de taxas específicas, as diferenças entre os valores de sobrevivência observados e o sistema SP-2 são muito pequenas para todas as idades, principalmente para o sexo feminino, como pode ser visto nos Gráficos 13 e 14. Esta concordância confirma que o novo sistema adaptativo de dois parâmetros SP-2 representa consistentemente a na-

tureza do padrão de mortalidade no tempo para o Estado de São Paulo.

Os Gráficos 15 e 16 mostram a comparação dos ajustamentos com os valores observados em termos do logaritmo da probabilidade de morrer (cuja escala procura acentuar as diferenças), ilustrado para o sexo feminino, respectivamente para 1970 e 1950. O modelo SP-2 mostra um valor mínimo que é localizado no grupo etário de 10-14 anos, como revelam os dados originais. A principal característica do modelo reside no *plateau* formado nas idades entre 15 e 25 anos – isto é, há uma descida relativamente suave até o ponto mínimo e uma subida depois deste ponto. Como pode ser visualizado, este *plateau* não se mostrou evidente para o ano de 1970, e sim para o ano de 1950. Tal pode ser devido a uma mudança no padrão de mortalidade para estes grupos etários durante o período 1950-70 no Estado de São Pau-

Gráfico 13
Diferenças entre $\ell(x)$ Obs./SP-2
1940-80 – Masculino

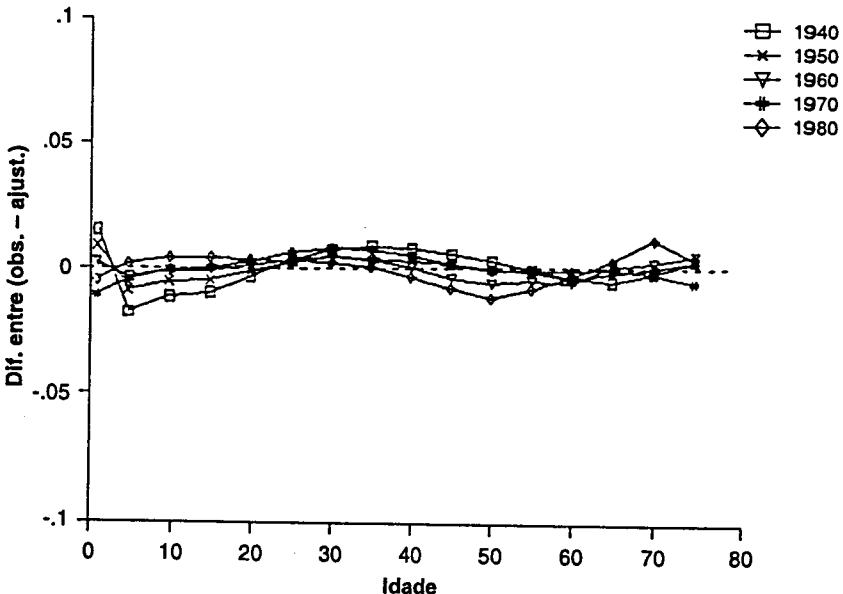


Gráfico 14
Diferenças entre $t(x)$ Obs/SP-2
1940-80 - Feminino

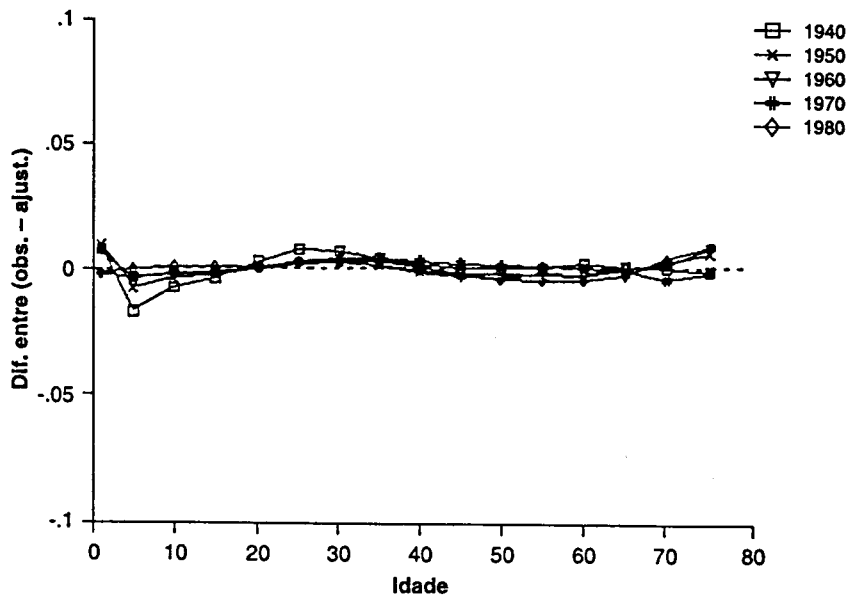


Gráfico 15
Ajuste do modelo SP-2
1970 - Feminino

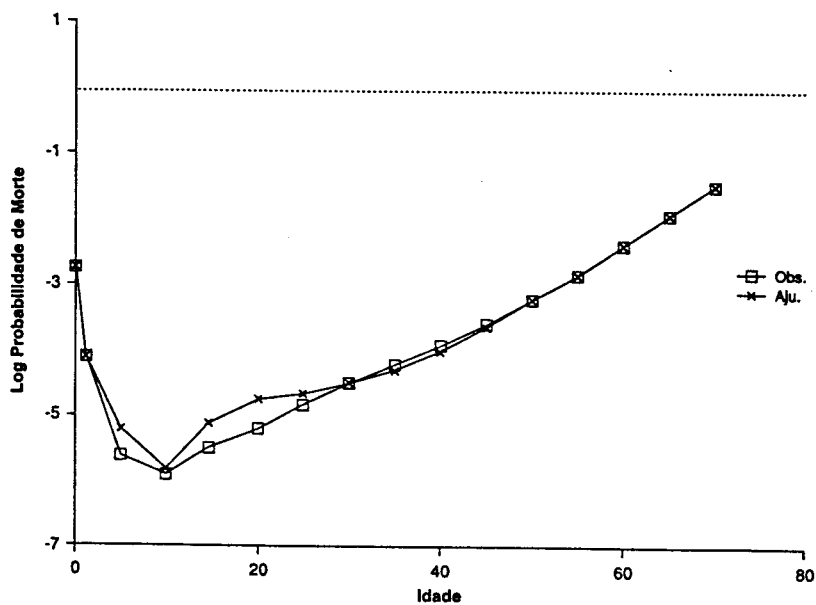
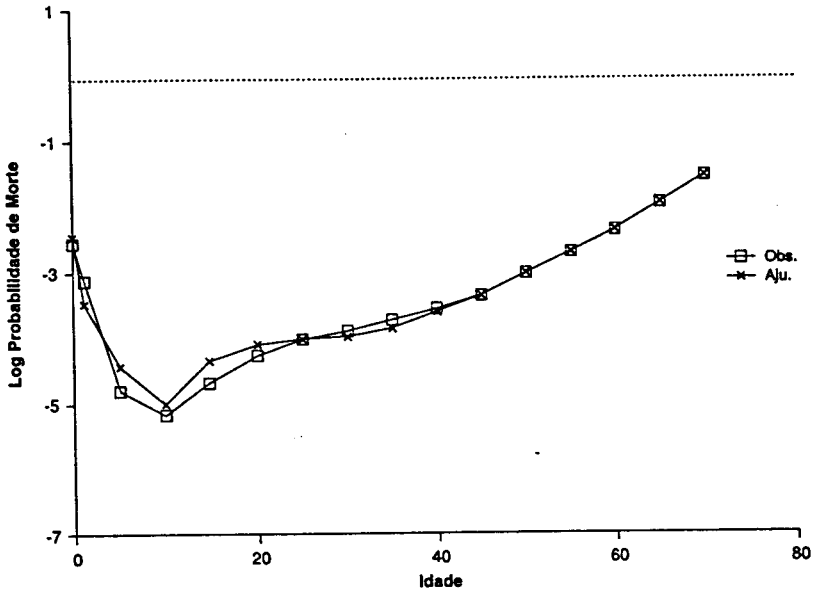


Gráfico 16
Ajuste do modelo SP-2
1950 - Feminino



lo, cujo modelo se ajustou melhor para estes grupos no passado mais remoto.

Uso do sistema de tábuas de vida SP-2

O sistema logito de quatro parâmetros pode ser aplicado para qualquer padrão de mortalidade sempre que se necessite de variações do modelo padrão de uma maneira flexível. Existem vários propósitos para os quais os modelos de tábuas de vida são usados para modelar e ajustar dados observados, tais como: suavizar um conjunto de dados para eliminar as irregularidades nas taxas de mortalidade, aproximando-o, desta forma, dos padrões de mortalidade típicos; avaliar a qualidade dos dados; construir tábuas de vida completas quando se dispõe de dados parciais; substituir dados reais quando estes são totalmente desconhecidos; projetar taxas de mortalidade.

Os modelos de quatro parâmetros abrem um novo caminho para explorar a composição das famílias de tábuas de vida e fornecer refinamentos de típicos padrões de mortalidade. Também podem ser usados para categorizar conjuntos de tábuas de vidas, mediante o exame de detalhadas diferenças geográficas entre regiões ou séries históricas pertencentes a uma mesma população.

O ajuste ou modelagem destes sistemas para dados observados ou tábuas de vida, contudo, pressupõe um conhecimento detalhado do padrão de mortalidade, particularmente nas idades extremas. Seu uso é supérfluo se a mortalidade na infância ou em idades avançadas não é conhecida. Mas, em certas aplicações, dados os valores de um dos parâmetros, o(s) outro(s) pode(m) ser identificado(s) ou eliminado(s). Uma vez que o padrão da mortalidade é identificado, esta informação pode ser usada para reduzir o

modelo a uma forma mais simples, com dois ou três parâmetros.

As primeiras tentativas de estender o uso do sistema de tábuas de vida SP-2 para representar e modelar a distribuição de mortalidade observada em várias áreas urbanas (capitais) do Brasil em 1970 e 1980 têm sido satisfatórias, demonstrando a sua eficácia para lidar com diferentes tipos de mortalidade. Possivelmente este sistema tem uma aplicação mais ampla. Tentativas de usar o SP-2 para dados de estados brasileiros produziram ajustes satisfatórios somente para as idades adultas (Paes, 1993). Ainda não foi possível encontrar um único modelo que seja representativo da mortalidade para qualquer população, de modo que um modelo que funciona para uma certa região pode não servir para outras.

Pode-se dizer que, historicamente, o modelo de tábuas de vida ajustado para São Paulo representa, consistentemente, o seu padrão de mortalidade experimentado no período 1940-80 para ambos os sexos. Abre-se, assim, um caminho exploratório com respeito à qualidade das estatísticas vitais no Brasil e à construção de modelos de tábuas de vida representativos da mortalidade das regiões brasileiras, caminho que deve ser seguido no futuro.

Apêndice:

Modelo adaptativo de quatro parâmetros de Ewbank

Dois sistemas de quatro parâmetros foram desenvolvidos até o momento. O primeiro foi formulado por Zaba (1979) e o segundo por Ewbank *et al.* (1983), aperfeiçoando o modelo anterior. Estes sistemas foram criados para situações em que os modelos de dois parâmetros não fossem suficientes para ajustar padrões de mortalidade observados.

O sistema de quatro parâmetros de Ewbank é uma extensão do sistema relacional de dois parâmetros de Brass. Os autores desenvolveram um modelo de quatro parâmetros, usando métodos de regressão linear, no qual o terceiro e quarto parâmetros atuam independentemente nas idades jovens e idosas, respectivamente. Neste modelo, se $\ell_s(x)$ é um conjunto de valores de sobrevivência na idade x , de qualquer modelo padrão escolhido, e κ e λ são os dois parâmetros extras, o modelo é definido por:

$$Y(x) = \alpha + \beta T[\ell_s(x); \kappa, \lambda]$$

onde

$$\frac{\left[\frac{\ell_s(x)}{1 - \ell_s(x)} \right]^\kappa - 1}{2\kappa} \quad \text{para } \ell_s(x) \geq 0,5$$

$$T[\ell_s(x); \kappa, \lambda] =$$

$$1 - \frac{\left[\frac{1 - \ell_s(x)}{\ell_s(x)} \right]^\lambda}{2\lambda} \quad \text{para } \ell_s(x) < 0,5$$

Este modelo converge para o modelo logito quando $\kappa = \lambda = 0$. Os quatro parâmetros do modelo são α , β , κ e λ . Os efeitos de α e β são os mesmos, como no sistema logito de dois parâmetros de Brass. Dado que κ é usado somente na transformação para os valores de $\ell_s(x)$ maiores que ou iguais a 0,5, ele afeta somente a mortalidade abaixo dos 51 anos. Similarmente, λ afeta somente a mortalidade nas idades acima dos 51 anos. Os efeitos desses parâmetros em termos de $\ell(x)$ podem ser visualizados em Ewbank *et al.* (1983).

Referências bibliográficas

- ALTMAN, A.M.G. e FERREIRA, C. E. "Evolução do censo demográfico e registro civil como fonte de dados para a análise da fecundidade e mortalidade no Brasil". *Revista Brasileira de Estatística*, Rio de Janeiro, 40(160), out.-dez., 1979, pp. 329-43.
- BRASS, W. "Methods for estimating fertility and mortality from limited and defective data". *Occasional Publication*, Chapel Hill, International Program of Laboratories for Population Statistics, 1975.
- COALE, A. J. e DEMENY, P. *Regional model life tables and stable population*. New Jersey, Princeton University Press, 1966.
- EWBANK, D.C.; LEON, J.C.G. e STOTO, M.A. "A reducible four-parameter system of model life tables". *Population Studies*, Nova York, United Nations, n. 37, 1983, pp.105-27.
- FERREIRA, C.E.C. "Tábuas abreviadas de mortalidade para o Estado de São Paulo – 1939/41, 1949/51, 1959/61 e 1969/71". *Informe Demográfico*, São Paulo, Fundação SEADE, vol. 4, 1980, pp.1-47.
- FRIAS, L.A.M. "Tábuas de mortalidade de alguns municípios das capitais e a influência dos óbitos de não residentes nessas capitais". *Revista Brasileira de Estatística*, Rio de Janeiro, 35(137), jan.-mar., 1974.
- FRIAS, L.A.M. e RODRIGUES, P. "Brasil: tábuas-modelo de mortalidade e populações estáveis". *Estudos e Pesquisas*, Rio de Janeiro, IBGE, vol.10, 1981.
- IBGE. *Recenseamento Geral do Brasil – 1940, 1950, 1960, 1970 e 1980 (Brasil e Estados)*. Rio de Janeiro, IBGE.
- IBGE, Centro Brasileiro de Estudos Demográficos. "Utilização das tábuas de vida modelo para se estimar a vida média do Brasil". *Revista Brasileira de Estatística*, Rio de Janeiro, 35(137), jan.-mar., 1974, pp. 29-34.
- JOHNSTON, J. *Econometric methods*. 3a. ed., Cingapura, McGraw-Hill, Inc., 1987.
- NAÇÕES UNIDAS. "Model life tables for developing countries". *Population Studies*, Nova York, United Nations, n. 77, 1982.
- PAES, N.A. *Model life table representation for Brazilian mortality*. PhD. Thesis, Faculty of Medicine, London School of Hygiene and Tropical Medicine, University of London, 1993.
- PUFFER, R.R. e SERRANO, C.V. "Patterns of mortality in childhood". *Pan American Health Organization*, Washington, D.C., 1973.
- ZABA, B. "The four-parameter logit life table system". *Population Studies*, Nova York, United Nations, n. 33, 1979, pp. 79-100.

RESUMO – Geração de um modelo de tábuas de vida representativo da mortalidade do Estado de São Paulo. O artigo investiga a possibilidade de se construir um conjunto de tábuas de vida modelo para o Estado de São Paulo, uma vez que se encontra disponível para o estado a mais longa série de registros de óbitos do país (1940-1980), praticamente completos. Propõe um novo modelo de tábuas de vida de dois parâmetros para São Paulo, designado de SP-2. Para ajustar os dados de sobrevivência, são testados os seguintes modelos: o modelo padrão geral de Brass, o modelo Brasil, o sistema de quatro parâmetros de Zaba e o sistema adaptativo de quatro parâmetros de Ewbank et al. Os resultados mostraram que o sistema de dois parâmetros SP-2 representa satisfatoriamente o padrão de mortalidade do Estado de São Paulo. Abrem-se, assim, novas possibilidades para a investigação da qualidade dos registros de óbitos e para a construção de modelos de tábuas de vida para outras regiões do país.

ABSTRACT – Generation of a model life table representation for mortality of the State of São Paulo. The possibility of constructing a particular set of model life table for

São Paulo is investigated in this paper since the most long set of registered death data in the country (from 1940 to 1980) is available for this state, and its quality has been proved to be practically complete. A new model life table, SP-2, is proposed. To fit the survivorship data for São Paulo the following models are tested here: Brass' general standard model, the Brazil's model, the Zaba four-parameter system and Ewbank et al.'s four adaptive system. The results showed that the simple São Paulo two-parameter system, SP-2, represents satisfactorily the mortality pattern of this state. It opens the way to continued investigations on the quality of the death registration data and improvement in the construction of model life tables for other regions.

(Recebido para publicação em julho de 1995)