MAT02262 - Estatística Demográfica I

Tábuas de vida: a estrutura

Rodrigo Citton P. dos Reis
citton.padilha@ufrgs.br

Universidade Federal do Rio Grande do Sul
Instituto de Matemática e Estatística
Departamento de Estatística

Porto Alegre, 2024

Instituto de MATEMÁTICA E ESTATÍSTICA UFRGS

Relembrando

Relembrando

Tábua de Vida de uma Geração vs. Tábua de Vida de Coorte Sintética

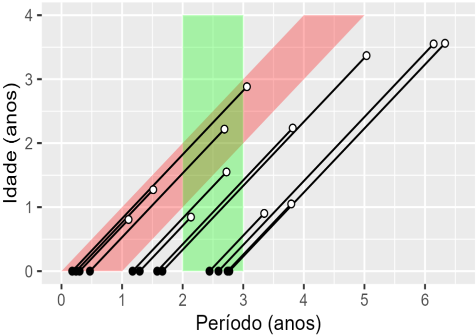


Figura 1: Diagrama de Lexis.

Tábua de Vida de Coorte Sintética

Esta tábua considera a experiência de mortalidade de uma dada população, num período curto de tempo (ano), e projeta a duração de vida, de cada indivíduo, baseada nas probabilidades reais de morte, numa **coorte hipotética** de nascidos vivos.

- Há, então, um padrão fictício de condições de mortalidade, dado que nenhuma coorte realmente experimentou ou experimentará este modelo particular de mortalidade.

Tábua de Vida de Coorte Sintética

- Este tipo de tábua responde também às indagações mencionadas anteriormente, levando-se em conta as seguintes pressuposições:
  - a **mortalidade**, em cada idade, mantém-se **constante** e igual à do ano-calendário, no qual a tábua é baseada;
  - a **população** exposta é **estacionária**, isto é, o número anual de nascidos vivos é igual ao número de mortes; o saldo migratório é nulo, ano após ano.

Formas de apresentação

Quanto à apresentação, pode-se ter **tábua de vida completa** ou **tábua de vida abreviada**.

- A diferença está no tamanho dos grupos etários considerados.
  - Na **tábua completa**, os grupos etários representam **um ano**, enquanto nas **abreviadas** ter-se-ão grupos de **cinco ou dez anos** de idade.

Análise de sobrevivência - conceitos básicos

- A variável aleatória não negativa T, que representa o tempo de falha, é usualmente especificada em análise de sobrevivência pela sua **função de sobrevivência** ou pela **função de taxa de falha (ou risco)**.

Função de sobrevivência

- A **função de sobrevivência** é definida como a **probabilidade de uma observação não falhar até um certo tempo** t, ou seja, a **probabilidade de uma observação sobreviver ao tempo** t.
- Em termos probabilísticos, isto é escrito como

$$S(t) = \Pr(T > t).$$

- Logo, a função de distribuição de probabilidade pode ser expressa como $F(t) = 1 - S(t)$ (probabilidade de uma observação não sobreviver ao tempo t).

Função de sobrevivência

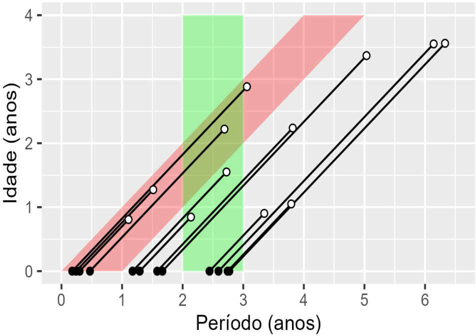
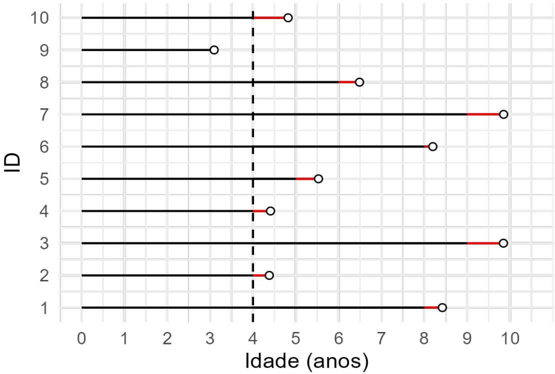
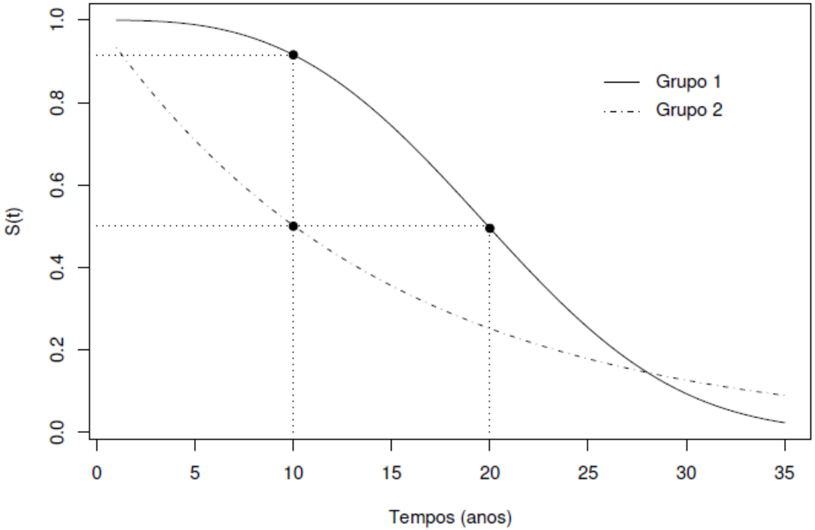


Figura 2: Diagrama de Lexis.

Função de sobrevivência



Função de sobrevivência



Função de sobrevivência

- A probabilidade da falha ocorrer em um intervalo de tempo $[t_1, t_2)$ pode ser expressa em termos da função de sobrevivência como

$$\Pr[T \in [t_1, t_2)] = \Pr(t_1 \leq T < t_2)] = S(t_1) - S(t_2).$$

Função de sobrevivência

- A probabilidade da falha ocorrer em um intervalo de tempo $[t_1, t_2)$ dado que a falha não ocorreu até o tempo t_1 também pode ser expressa em termos da função de sobrevivência:

$$\Pr(t_1 \leq T < t_2 | T > t_1) = \frac{\Pr(t_1 \leq T < t_2)}{\Pr(T > t_1)} = \frac{S(t_1) - S(t_2)}{S(t_1)}.$$

Tábuas de vida: a estrutura

Considerações gerais

Até o momento, os coeficientes específicos de mortalidade foram considerados como uma das melhores **medidas de mortalidade**, pois estimam o **risco de morrer**.

- No entanto, esses coeficientes não respondem à indagação:

Qual a probabilidade

de uma pessoa com idade exata x, no início de um determinado ano, vir a falecer neste mesmo ano?

Considerações gerais

Para a estimativa dessa probabilidade, há necessidade de lembrar que a probabilidade de ocorrência de um dado evento E é

- a relação entre o número de casos que compõem o evento E e o número total de casos possíveis, desde que cada caso possível seja igualmente provável de ocorrer.

Considerações gerais

Para o cálculo da **probabilidade de morte (na idade x)** dos indivíduos de diversas idades (q_x), há necessidade de se conhecer o total de casos possíveis (E_x) e o total de casos que compõem o evento E (D_x).

Considerações gerais

Para tanto, há de ser considerado:

- D_x: número de casos que compõem o evento **"vir a falecer durante o período"**;
  - ou seja, aproximadamente, o **número de óbitos**, ocorridos naquele período, das pessoas com idade x no início do período.
- E_x: número de casos possíveis;
  - ou seja, número de indivíduos com exata idade x, no início do período, isto é, **expostos ao risco de morrer**.

Considerações gerais

Para o conhecimento destas duas quantidades, há de se refletir sobre os seguintes fatos:

- Os **dados oficiais** mostram, para cada ano calendário, o número de óbitos registrados para cada idade x.
- Entretanto, D_x deve representar os óbitos das pessoas que tinham idade x no início do período (ano), e não sua idade ao morrer.

Considerações gerais

- Assim, se uma pessoa completar 20 anos (x anos) em março e falecer em setembro, seu óbito será registrado e classificado como sendo de uma pessoa com 20 anos (x anos).
- Da mesma maneira, ocorre que pessoas que tenham 20 anos (x anos) no início do período venham a completar 21 anos ($x+1$ ano) em março, vindo a falecer em setembro.
  - O óbito, também aqui, será registrado como o de uma pessoa com $x+1$ anos (21 anos): não será levado em conta nos cálculos de óbitos de pessoas com x anos (20 anos).

Considerações gerais

- Portanto, ao se usar os dados oficiais no cálculo das probabilidades de morrer, por exemplo, aos 20 anos, ter-se-á um caso a mais no número de óbitos de pessoas que faleceram aos 20 anos (x anos), tendo $x-1$ anos (19 anos) no início do referido período;
  - e um a menos, no número de óbitos de pessoas que tinham 20 anos (x anos) no início do período, mas vieram a falecer com 21 anos ($x+1$ anos) no período considerado.

A pressuposição necessária é que estes dois fatos **se compensem** e, portanto, o valor D_x seja, em média, o mesmo valor fornecido pelo Registro Civil.

Considerações gerais

- A **população exposta ao risco** (E_x) deverá ser composta por indivíduos cujas idades, em média, sejam iguais a x anos, no início do período.
  - Para tanto, tem-se, como fonte de informação, a **população recenseada** e estimada para a **metade do ano (1º de julho)**.
- Essa população terá, em média, $x + 0,5$ anos (P_x) e corresponderá à população que terá, em média, x anos no **início do período (1º de janeiro)**.

Diz-se *corresponderá* e não *será igual*, pois há indivíduos com exata idade x, no começo do ano, qua falecem no primeiro semestre e não serão, portanto, recenseados.

Considerações gerais

Para se estimar E_x, há necessidade da seguinte aproximação:

$$E_x = P_x + \text{ óbitos de idade } x \text{ ocorridos na primeira metade do período.}$$

Considerações gerais

Para se conhecer o número de óbitos ocorridos na primeira metade, basta pressupor que os óbitos estejam **igualmente *(uniformemente)* distribuídos** durante o período, isto é, que **a metade dos óbitos ocorra na primeira metade do período**.

- Esta pressuposição é válida somente para os óbitos de indivíduos com idade igual, ou superior, a dois anos.
- Nas idades abaixo de dois anos, sabe-se que grande parte dos óbitos ocorre no início da vida.

Considerações gerais

A pressuposição anteriormente mencionada não poderá ser aplicada.

- É uso corrente adotar-se a pressuposição de que, em menores de um ano, 2/3 (0,667) dos óbitos ocorrem na primeira metade do período e que, dos óbitos de crianças de um ano completo de idade, 3/5 (0,60) ocorrem na primeira metade.
- **Oya**[1], observando tal fato para o Município de São Paulo, em 1968/1969, recomenda que estas proporções sejam, para menores de um ano, iguais a 0,863 e 0,826, respectivamente, nos sexos masculino e feminino, e, para os óbitos de um ano, seja igual a 0,70.

[1] OYA, D. R. T. *Estudo da distribuição do fator de separação* f_x^t *na tábua de sobrevivência*. São Paulo, 1970. Dissertação de Mestrado — Faculdade de Saúde Pública da USP.

Considerações gerais

Assim, para o cálculo da probabilidade de morrer (q_x) adotam-se as seguintes estimativas:

a) para indivíduos com **idade igual ou superior a dois anos**:

$$q_x = \frac{D_x}{P_x + \frac{D_x}{2}}.$$

Considerações gerais

b) para indivíduos com **idade inferior a dois anos**:

b1) menores de um ano, **sexo masculino**

$$q_0 = \frac{D_0}{P_0 + 0,863 D_0};$$

b2) menores de um ano, **sexo feminino**

$$q_0 = \frac{D_0}{P_0 + 0,826 D_0};$$

Considerações gerais

b3) indivíduos com **um ano de idade**

$$q_1 = \frac{D_1}{P_1 + 0,70D_1}.$$

Considerações gerais

- Quando se constrói uma tábua de sobrevivência do **tipo abreviada**, isto é, os grupos etários são de tamanho igual a cinco ou dez anos, há necessidade de modificações, pois o que se pretende é estimar a probabilidade de um indivíduo do grupo etário $x \vdash x + n$ vir a falecer, durante o período que estaria nesse grupo (período de n anos).

Considerações gerais

- Portanto, deve-se multiplicar o número de óbitos (D_x) por n.
- Tal probabilidade será, então:

$$ {}_nq_x = \frac{{}_nO_x}{{}_nP_x + \frac{{}_nO_X}{2}}, $$

em que:

$$ {}_nO_x = n \times {}_nD_x. $$

- As estimativas das probabilidades de morrer, para os residentes masculinos no Município de São Paulo, em 1970, estão apresentadas na tabela a seguir.

Considerações gerais

Tabela 8.1. **Probabilidade de morrer, segundo grupos etários, população masculina do Município de São Paulo, 1970.**

Grupos etários (em anos)	População em 1/7/1970 $_nP_x$	Óbitos $_nD_x$	Óbitos esperados $_nO_x$	Expostos a morrer $_nE_x$	Probabilidade de morrer $_nq_x$
0	66 163	6 578	6 578	71 840	0,09157
1	60 223	338	338	60 460	0,00559
2	65 101	183	183	65 192	0,00281
3	66 889	117	117	66 948	0,00175
4	68 392	79	79	68 432	0,00115
5 ⊢ 10	334 692	256	1 280	335 332	0,00382
10 ⊢ 15	300 761	222	1 110	301 316	0,00368
15 ⊢ 20	275 881	383	1 915	276 839	0,00692
20 ⊢ 25	298 175	600	3 000	299 675	0,01001
25 ⊢ 30	256 425	708	3 540	258 195	0,01371
30 ⊢ 35	225 343	851	4 255	227 471	0,01871
35 ⊢ 40	199 436	996	4 980	201 926	0,02466
40 ⊢ 45	182 523	1 240	6 200	185 623	0,03340
45 ⊢ 50	138 689	1 320	6 600	141 989	0,04648
50 ⊢ 55	105 660	1 504	7 520	109 420	0,06873
55 ⊢ 60	83 668	1 662	8 310	87 823	0,09462
60 ⊢ 65	64 066	1 961	9 805	68 969	0,14217
65 ⊢ 70	43 796	1 971	9 855	48 724	0,20226
70 ⊢ 75	27 285	1 902	9 510	32 040	0,29682
75 ⊢ 80	13 211	1 360	6 800	16 611	0,40937
80 ⊢ 85	6 766	873	4 365	8 949	0,48779
85 e +	4 717	650	4 717	4 717	1,00000

Considerações gerais

Alguns exemplos dos cálculos são mostrados a seguir:

$$
\begin{aligned}
q_0 &= \frac{6578}{66163 + (0,863 \times 6578)} = 0,091565; \\
q_1 &= \frac{338}{60223 + (0,70 \times 338)} = 0,005591; \\
q_2 &= \frac{183}{65.101 + (1/2 \times 183)} = 0,002807;
\end{aligned}
$$

Considerações gerais

$$ {}_5q_5 = \frac{5 \times 256}{334692 + (1/2 \times 5 \times 256)} = 0,003817. $$

- A probabilidade de morrer, no **último grupo etário**, é igual a $1,0$, pois a coorte se extingue.

Comentários finais

1. Note que ao multiplicar o numerador e o denominador de q_x por $1/P_x$, temos que

$$
\begin{aligned}
q_x &= \frac{D_x/P_x}{P_x/P_x + \frac{D_x/P_x}{2}} \\
&= \frac{m_x}{1 + \frac{m_x}{2}},
\end{aligned}
$$

em que m_x são as taxas de mortalidade específicas por idade.

Comentários finais

2. O valor 0,5 que multiplica D_x é um **fator de separação** 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.

Próxima aula

- Tábuas de vida (continuação).

Para casa

- Ler o capítulo 9 do livro "Métodos Demográficos Uma Visão Desde os Países de Língua Portuguesa"[2].

[2] FOZ, Grupo de. *Métodos Demográficos Uma Visão Desde os Países de Língua Portuguesa*. São Paulo: Blucher, 2021. https://www.blucher.com.br/metodos-demograficos-uma-visao-desde-os-paises-de-lingua-portuguesa_9786555500837

Por hoje é só!

Bons estudos!