

Modelos de Análise de Sobrevivência

**Ernesto F. L. Amaral
Magna M. Inácio**

**02 de dezembro de 2010
Tópicos Especiais em Teoria e Análise Política:
Problema de Desenho e Análise Empírica (DCP 859B4)**

TAXA \neq PROBABILIDADE

- Antes de tudo, é importante entender a diferença entre taxa e probabilidade.
- Taxa pode ser definida como a divisão do número de eventos, que ocorreram em um determinado intervalo de tempo, pelo número de indivíduos que estiveram expostos ao risco do evento em questão, durante o mesmo período.
- A probabilidade é similar à taxa, com a diferença de que o denominador é composto por todas as pessoas de uma determinada população no começo do período de observação.

QUAL É ENTÃO O DENOMINADOR DA TAXA? (Carvalho, Sawyer e Rodrigues, 1998)

- Para o cálculo de taxas, uma vez que a população total muda a cada instante no decorrer do tempo, surge a dúvida sobre qual população levar em consideração no denominador.
- No início do período, não estão incluídos os indivíduos que entraram na população durante o tempo considerado.
- As pessoas que estão presentes no começo do período e que saem antes do fim do período não podem ter o mesmo peso que aquelas que permanecerão (sobreviverão) durante todo o tempo considerado.

O CONCEITO DE PESSOAS-ANO

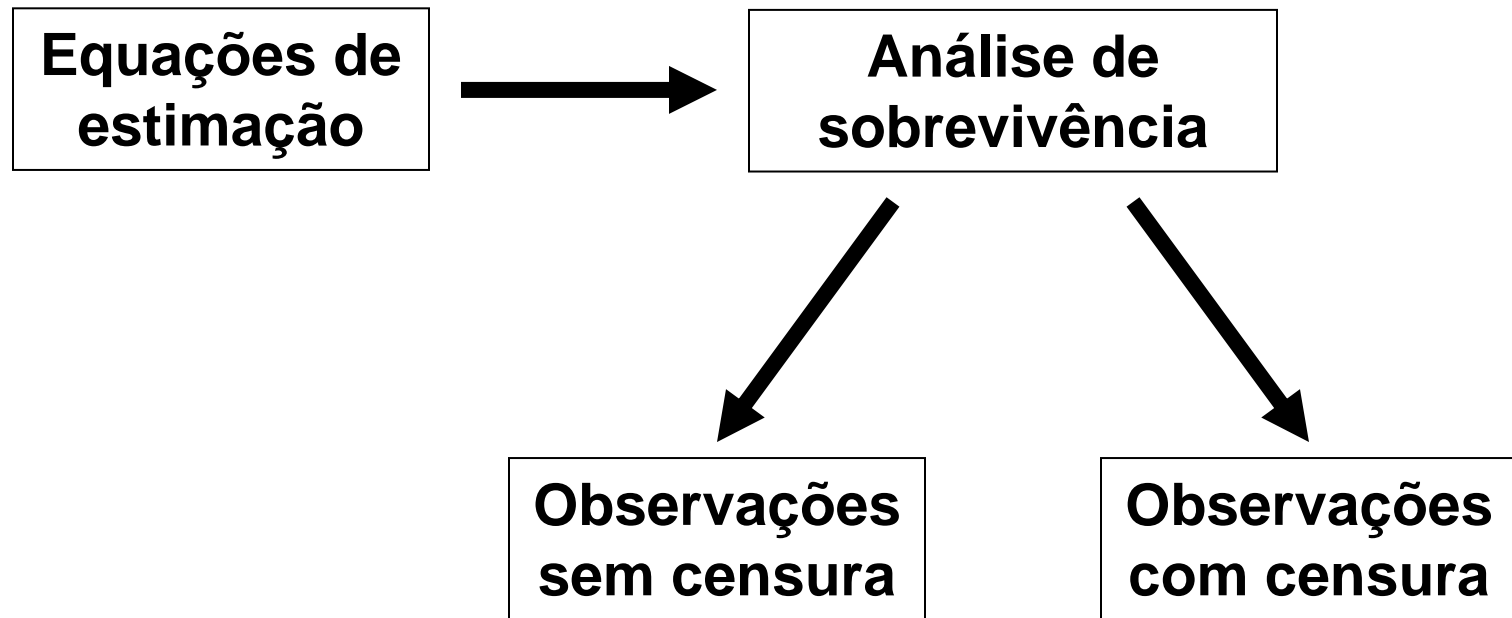
- O ideal seria contabilizar no denominador o número de pessoas-ano da população em estudo.
- Todo indivíduo presente do início ao fim do ano (período de tempo) deveria ser contado como uma pessoa-ano.
- Indivíduos que estavam no início do período e que não permanecem durante todo tempo, assim como os que são inseridos durante o período, deveriam ser contabilizados pela fração de tempo que estiveram presentes.
- O conceito de pessoas-ano é ideal para especificar a população exposta ao risco de um evento, sendo o denominador ideal para taxas.
- É muito difícil calcular o número de pessoas-ano.

QUAL A SOLUÇÃO?

- O denominador das taxas deveria ser o produto do número de pessoas multiplicado pelo tempo (anos, meses, dias...), ou frações de tempo, que cada pessoa esteve exposta em um local e período.
- Em substituição ao número de pessoas-ano, o denominador das taxas pode utilizar a população total no meio do período, na suposição de que a inclusão e exclusão de indivíduos na população ocorreram uniformemente no decorrer desse período.
- Para análise de sobrevivência, é importante obter o dado (variável de exposição) que indica o momento de entrada e saída de cada observação (indivíduo, instituição, país...) na situação em análise (variável de interesse).

DADOS CENSURADOS

- Censura implica que sistema pode falhar mesmo antes que todos itens tenham falhado em um determinado tempo.
- Este fato tem determinadas razões.



ANÁLISE DE SOBREVIVÊNCIA

- A análise de sobrevivência envolve a modelagem do tempo da ocorrência de um evento, onde a interrupção é considerado um evento.
- As técnicas estatísticas conhecidas como análise de sobrevivência são utilizadas quando se pretende analisar um fenômeno, em relação a um **período de tempo**.
- É analisado o tempo transcorrido entre um **evento inicial** (no qual um sujeito estava em um estado particular) e um **evento final** (que modifica este estado).

MÉTODOS ESTATÍSTICOS NA ANÁLISE DE SOBREVIVÊNCIA

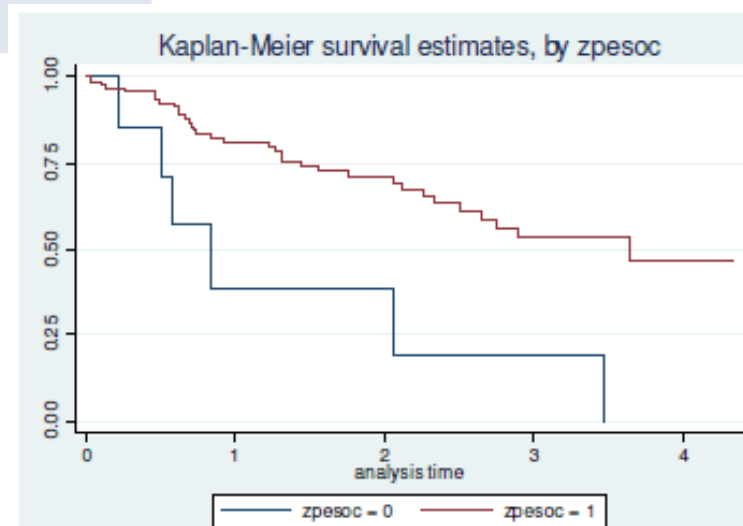
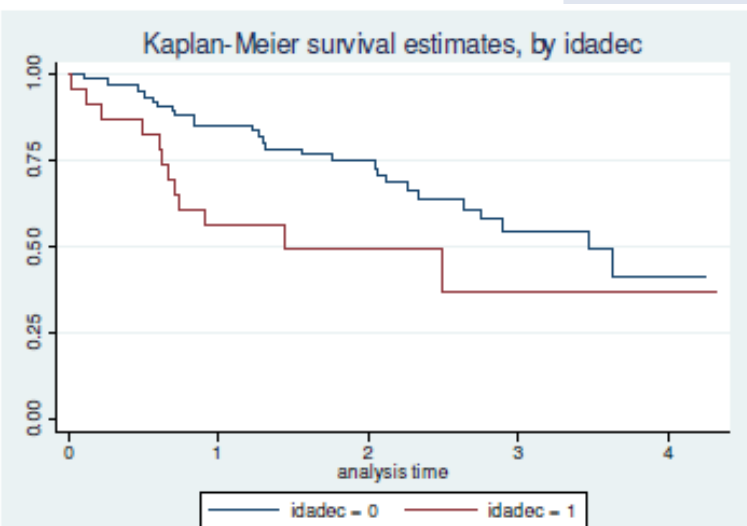
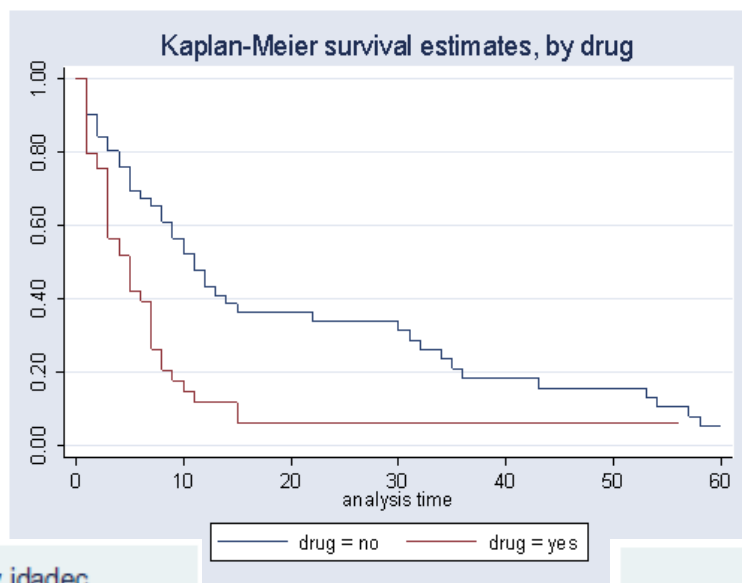
- Uma série de métodos estatísticos podem ser utilizados na análise de sobrevivência:
 - Estimativas de sobrevivência de Kaplan-Meier.
 - Teste estatístico de Log-Rank.
 - Modelo de risco proporcional de Cox.

COMPARAÇÃO DE CURVAS DE SOBREVIVÊNCIA

- Freqüentemente, desejamos comparar curvas de sobrevivência de diferentes grupos.
- Devemos dividir o conjunto total de observações em grupos distintos.
- Em seguida, são estimadas as funções de sobrevivência separadamente para cada grupo.
- É ainda possível utilizar o teste estatístico de Log-Rank:
 - H_0 : não há diferença entre as curvas.
 - H_1 : há diferença entre as curvas.
- Ou seja, curvas de sobrevivência de Kaplan-Meier podem ser estratificadas por subgrupos e comparadas pela estatística de Log-Rank.

ESTIMATIVAS DE SOBREVIVÊNCIA DE KAPLAN-MEIER ¹⁰

– Método de gráfico descritivo para comparação de sobrevivência entre diferentes grupos.



TESTE DE LOG-RANK NO STATA

```
sts test grupo
```

```
failure _d: Status == 1
analysis time _t: Tempo
```

```
Log-rank test for equality of survivor functions
```

	Events	Events
grupo	observed	expected
1	4	3.46
2	3	2.44
3	5	3.44
4	2	3.57
5	3	4.09
Total	17	17.00

```
chi2(4) = 2.17
```

```
Pr>chi2 = 0.7040
```

- Valor de “p” maior que 0,05 indica que não existe diferença estatisticamente significativa entre as curvas de sobrevivência, com nível de significância de 95%.

MODELO DE RISCO PROPORCIONAL DE COX

- Estudos em Ciência Política muitas vezes envolvem covariáveis que podem estar relacionadas com o tempo de sobrevivência.
- O modelo de Cox permite a análise de dados provenientes de estudos de “tempo de sobrevivência”, em que a resposta é o tempo até ocorrência de um evento de interesse, controlado por covariáveis.
- O modelo de Cox é caracterizado pelos coeficientes β 's que medem os efeitos das covariáveis sobre a função de taxa de falha.
- Este modelo estima razões de riscos (*hazard ratios*).

- Modelo de Cox: $\lambda(t) = \lambda_0(t) \exp\{\beta_1 X_1 + \dots + \beta_p X_p\}$, sendo:
 - $\lambda_0(t)$ = função de risco básica
 - β_1, \dots, β_p = parâmetros do modelo
 - X_1, \dots, X_p = variáveis explicativas
- No Stata:
 - Definindo os dados de sobrevivência (stset):
`stset months, failure (termination==1)`
 - Definindo o modelo de Cox (stcox):
`stcox mindur effpart minority surplus mwc coalition...`
- No exemplo, coeficientes com $p < 0,05$ indicam variáveis que são importantes para explicar duração do ministério.

CABINET TURNOVER IN PARLIAMENTARY DEMOCRACIES (Huber e Martinez-Gallardo, 2003 APSA)

Riscos competitivos

Término

Table 2. Determinants of Ministerial Duration, Competing Risks[†]

		Dependent Variable: Ministerial Duration, in Months (Competing Risks)			
Independent Variables:		(1) Pooled	(2) Terminations	(3) Reshuffles	
Cabinet Duration	Effective Parties	1.061 (2.48)	1.070 (2.69)	1.020 (0.28)	
	Minority	1.303 (3.48)	1.274 (2.91)	1.546 (2.24)	
	Surplus	0.901 (-1.07)	0.870 (-1.32)	1.117 (0.43)	
	MWC	0.594 (-6.76)	0.563 (-6.8)	0.930 (-0.37)	
	Government Heterogeneity	1.021 (1.77)	1.019 (1.44)	1.022 (0.75)	
	Electoral Volatility	1.092 (6.2)	1.107 (6.77)	1.001 (0.03)	
	Electoral Vol * ln(Time)	1.008 (2.58)	1.007 (2.05)	1.022 (1.79)	
	Investiture	0.909 (-1.51)	0.913 (-1.33)	0.884 (-0.74)	
	Uncertainty	Size of Minister's Party	0.980 (-7.93)	0.974 (-9.39)	1.012 (1.64)
		Assembly	1.000 (2.98)	1.000 (2.76)	1.000 (0.6)
Ln (Age of Democracy)		0.966 (-1.13)	0.943 (-1.78)	1.133 (1.3)	
Impact on Policy	Minister Autonomy	1.004 (0.08)	0.984 (-0.31)	1.122 (0.98)	
	Policy Value of Portfolio	1.221 (9.5)	1.195 (7.89)	1.368 (5.09)	
	Finance	0.915 (-1.57)	0.880 (-2.07)	1.215 (0.93)	
	Foreign Affairs	0.764 (-3.66)	0.741 (-3.73)	0.855 (-0.86)	
	Agriculture			0.500 (-2.83)	
Trust	Extreme Minister	1.047 (1.84)	1.064 (2.34)	0.944 (-0.81)	
Constraints	Last Election	0.990 (-23.78)	0.989 (-22.9)	1.000 (0.79)	
	Same Party as PM	1.490 (5.32)	1.611 (5.9)	0.989 (-0.06)	
	Minster Left-Right	0.987 (-0.99)	0.966 (-2.41)	1.100 (2.38)	
N		353682	353682	353682	
Ln (L)		-15853	-13370	-2435	

[†] Coefficients are hazard ratios. Z statistics are shown in parentheses.

** The hazard ratios for coalitions are .6312 (-6.09) for pooled hazards, .6005 (-6.16) for terminations, and .9723 (-0.15) for reshuffles.

Table 3. Determinants of Ministerial Duration, Terminations[†]

		Dependent Variable: Ministerial Duration, in Months (Terminations)		
Independent Variables:		(1)	(2)	(3)
Cabinet Duration	Effective Parties	1.148 (7.48)	1.091 (3.51)	1.070 (2.69)
	Minority	1.304 (4.05)	1.405 (4.17)	1.274 (2.91)
	Surplus		0.873 (-1.28)	0.870 (-1.32)
	MWC		0.572 (-6.62)	0.563 (-6.8)
	Coalition	0.832 (-2.59)		
	Government Heterogeneity	1.037 (3.47)	1.028 (2.20)	1.019 (1.44)
	Electoral Volatility	1.098 (6.53)	1.105 (6.69)	1.107 (6.77)
	Electoral Vol * ln(Time)	1.007 (2.29)	1.006 (1.86)	1.007 (2.05)
	Investiture	1.222 (4.22)	0.917 (-1.27)	0.913 (-1.33)
	Uncertainty	Size of Minister's Party		0.986 (-7.42)
Assembly			1.000 (3.22)	1.000 (2.76)
Ln (Age of Democracy)			0.930 (-2.22)	0.943 (-1.78)
Impact on Policy	Minister Autonomy		0.964 (-0.71)	0.984 (-0.31)
	Policy Value of Portfolio		1.170 (7.10)	1.195 (7.89)
	Finance		0.866 (-2.33)	0.880 (-2.07)
	Foreign Affairs		0.714 (-4.24)	0.741 (-3.73)
Trust	Extreme Minister			1.064 (2.34)
Constraints	Last Election	0.989 (24.87)	0.989 (23.7)	0.989 (-22.90)
	Same Party as PM			1.611 (5.9)
	Minster Left-Right			0.966 (-2.41)
N		371187	359201	353682
Ln (L)		-14913	-13516	-13370

[†] Coefficients are hazard ratios. Z statistics are shown in parentheses.

DURAÇÃO DO MINISTÉRIO EM MESES (TÉRMINO)

Table 3. Determinants of Ministerial Duration, Terminations[†]

Dependent Variable: Ministerial Duration, in Months (Terminations)

<i>Independent Variables:</i>		(1)	(2)	(3)
Cabinet Duration	Effective Parties	1.148 (7.48)	1.091 (3.51)	1.070 (2.69)
	Minority	1.304 (4.05)	1.405 (4.17)	1.274 (2.91)
	Surplus		0.873 (-1.28)	0.870 (-1.32)
	MWC		0.572 (-6.62)	0.563 (-6.8)
	Coalition	0.832 (-2.59)		
	Government Heterogeneity	1.037 (3.47)	1.028 (2.20)	1.019 (1.44)
	Electoral Volatility	1.098 (6.53)	1.105 (6.69)	1.107 (6.77)
	Electoral Vol * $\ln(\text{Time})$	1.007 (2.29)	1.006 (1.86)	1.007 (2.05)
	Investiture	1.222 (4.22)	0.917 (-1.27)	0.913 (-1.33)