

Análise de Sobrevivência – Modelos de Cura

Ana Maria Abreu, *abreu@uma.pt*

Universidade da Madeira, CCCEE, CCM

1. Introdução

A Análise de Sobrevivência consiste na análise de dados referentes a medições do tempo decorrido desde um instante inicial até à ocorrência de um determinado acontecimento, numa população. No entanto, nem sempre é possível observar o acontecimento de interesse devido, em geral, a certas restrições no processo de recolha dos dados, situação em que surgem as designadas observações censuradas.

Uma situação particular ocorre quando à maioria das observações censuradas correspondem os valores mais elevados. Este facto traduz-se graficamente na estabilização da curva da estimativa de Kaplan-Meier da função de sobrevivência (f.s.), como ilustra a Figura 1.

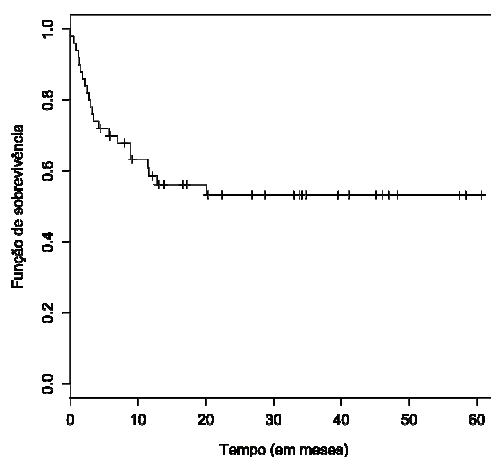


Figura 1: Dados de leucemia (Klein e Moeschberger, 1998, p.10)

Neste caso, suspeita-se da existência de indivíduos que nunca vão experimentar o acontecimento de interesse, ainda que o tempo em observação pudesse ser prolongado indefinidamente. Tais indivíduos são considerados imunes ao acontecimento de interesse e, nesta perspetiva, a função de sobrevivência que lhes corresponde, em qualquer instante, é sempre um (f.s. degenerada). Neste âmbito, a população em estudo é constituída por uma mistura de indivíduos imunes (ou curados ou não suscetíveis) e de indivíduos não imunes (ou doentes ou suscetíveis); a estes últimos corresponde uma f.s. não degenerada.

A observação do acontecimento de interesse permite que os indivíduos sejam classificados como suscetíveis. Quando esse acontecimento não é observado, em geral, não é possível identificar os indivíduos imunes, embora a sua presença seja admitida. Apenas no caso em que é definido algum critério de classificação para os indivíduos curados é que a identificação é inequívoca. Laska e Meisner (1992), Taylor (1995), Tamura *et al.* (2000) e Betensky e Schoenfeld (2001), fornecem dos poucos exemplos desta situação. Nos três primeiros, é definido um valor *a priori* e os indivíduos que tenham um tempo de vida superior a esse valor são considerados curados. Em Betensky e Schoenfeld (2001), a ocorrência de um certo acontecimento (alta hospitalar) distinto do acontecimento de interesse (morte pela doença) possibilita a identificação dos indivíduos curados.

Neste artigo, serão definidos os modelos de cura de mistura e de não-mistura. Será apresentado um modelo de cura de mistura baseado na distribuição de Chen (Chen, 2000, Abreu, 2005 e Abreu e Rocha, 2006) e ilustrada a estimação dos parâmetros através de dois exemplos de aplicação a dados reais.

2. Modelos de cura

Quando é admissível a existência de indivíduos curados na população, os modelos de cura, caracterizados por uma função de sobrevivência imprópria, constituem uma alternativa adequada aos modelos mais clássicos da Análise de Sobrevivência. Estes modelos dividem-se em dois grupos: os modelos de cura de mistura e os modelos de cura de não-mistura. Habitualmente designam-se apenas por modelos de cura, sempre que seja claro no contexto qual o tipo que está a ser considerado.

2.1 Modelo de cura de mistura

Sejam T uma variável aleatória (v.a.) que representa o tempo de vida de um indivíduo numa dada população e Y uma v.a. binária, onde $Y=1$ se o indivíduo é suscetível e $Y=0$ se o indivíduo é imune. Seja $S(t|Y=1)$ a função de sobrevivência (f.s.) do tempo de vida dos indivíduos suscetíveis e p a proporção de indivíduos imunes na população.

O modelo de cura definido à custa da f.s. é dado por

$$S(t) = p + (1 - p)S(t|Y = 1)$$

e à custa da função de risco é

$$h(t) = \frac{(1 - p)f(t|Y = 1)}{S(t)}.$$

Este caso particular dos modelos de mistura, designado por modelo de mistura não padrão, refere-se ao caso de uma mistura de duas distribuições, uma das quais degenerada. Note-se que $S(t)$ é uma f.s. imprópria pois $S(\infty)=p$, enquanto que $S(t|Y=1)$ é uma f.s. própria. Por outro lado, $h(t)$ tem integral finito em $(0, \infty)$, logo não se trata de uma função de risco no sentido usual.

Neste tipo de modelos de cura interessa estimar não só os parâmetros da função de distribuição dos indivíduos suscetíveis, como também a proporção de indivíduos imunes. Maller e Zhou (1992) propuseram uma estimativa não paramétrica desta proporção, que é obtida através do valor da estimativa de Kaplan-Meier da f.s. calculada na maior observação.

2.2 Modelo de cura de não-mistura

Uma vez que num modelo de cura a f.s. populacional é uma função imprópria e tendo em conta a relação existente entre a f.s. e a correspondente função de risco cumulativa, esta última vai ser uma

função limitada superiormente. Assim, seja $S(t)$ a f.s. de T e $H(t)$ é tal que $S(t)=\exp[-H(t)]$. Se $S(\infty)>0$, existe $\theta<\infty$, tal que $H(\infty)=\theta$. De acordo com Yakovlev *et al.* (1993), uma forma possível de caracterizar esta propriedade é considerar $H(t)=\theta F(t)$, onde $F(t)$ designa a função de distribuição (própria) de uma v.a. não negativa.

Deste modo, o modelo de cura de não-mistura pode ser escrito na forma

$$S(t)=\exp[-\theta F(t)],$$

sendo a correspondente função de risco dada por

$$h(t)=\theta f(t),$$

onde $f(t)$ é a função densidade correspondente a $F(t)$. Neste caso, $S(\infty)=e^{-\theta}$ corresponde à probabilidade de cura.

Neste contexto, existem duas hipóteses para obter uma f.s. imprópria. Ou se modifica o espaço de parâmetros de alguma distribuição própria como acontece, por exemplo, com a distribuição de Gompertz modificada (Cantor e Shuster, 1992), ou se considera uma variável fragilidade para caracterizar a suscetibilidade de cada indivíduo ao acontecimento de interesse. Nesta última situação, é a escolha da distribuição da fragilidade e não da distribuição do tempo de vida que vai dar origem a uma função de sobrevivência imprópria. Um exemplo deste caso é a distribuição Qui-quadrado não central com zero graus de liberdade (Rocha, 1995).

3. Modelo de cura baseado na distribuição de Chen

Esta secção é constituída por quatro partes. A primeira apresenta a distribuição de Chen, uma vez que ainda é algo desconhecida. Em seguida, surge o modelo de cura com base nesta distribuição, após o que é indicado a metodologia usada para a estimação dos parâmetros. Por último, é efetuada uma aplicação a dados reais.

3.1 Distribuição de Chen

A função de distribuição proposta por Chen (2000) é

$$F(t)=1-\exp\left[-\lambda\left(1-\exp\left(t^\beta\right)\right)\right], \quad t > 0,$$

onde λ e β são os parâmetros da distribuição, sendo λ o parâmetro de escala e β o parâmetro de forma. Assim, as correspondentes f.s. e função de risco são, respetivamente,

$$\bar{F}(t)=\exp\left[-\lambda\left(1-\exp\left(t^\beta\right)\right)\right], \quad t > 0 \quad \text{e} \quad h^*(t)=\lambda\beta t^{\beta-1}\exp\left(t^\beta\right) \quad t > 0.$$

A função $h^*(t)$ é crescente quando $\beta \geq 1$ e, para $\beta < 1$, é decrescente para $t \in \left[0, \left(\frac{1}{\beta}-1\right)^{\frac{1}{\beta}}\right]$ e é

crescente para $t \geq \left(\frac{1}{\beta}-1\right)^{\frac{1}{\beta}}$.

3.2 Modelo de cura baseado na distribuição de Chen

Considere-se a distribuição de Chen para a distribuição do tempo de vida dos indivíduos suscetíveis. O modelo de cura, definido à custa da f.s. é dado por

$$S(t) = p + (1-p)\exp\left[\lambda\left(1 - \exp\left(t^\beta\right)\right)\right], \quad t > 0.$$

Se o modelo for escrito à custa da função de risco, tem-se

$$h(t) = \frac{(1-p)\lambda\beta t^{\beta-1} \exp\left(t^\beta\right) \exp\left[\lambda\left(1 - \exp\left(t^\beta\right)\right)\right]}{p + (1-p)\exp\left[\lambda\left(1 - \exp\left(t^\beta\right)\right)\right]}.$$

Importa notar que, embora o parâmetro β continue a ser um parâmetro de forma do modelo populacional, o comportamento da função de risco populacional é completamente diferente do da função de risco dos indivíduos suscetíveis: $h(t)$ é decrescente para $\beta \leq 1$ e é unimodal para $\beta > 1$. Além disso, no caso em que $\beta > 1$, o valor da moda será tanto maior quanto menor for o valor de p , o que faz todo o sentido uma vez que p representa a proporção de indivíduos imunes. Relativamente ao parâmetro λ , de certa forma vai definir o valor de t a partir do qual a função de risco irá estabilizar num valor próximo de zero, ou seja, neste contexto, o valor de t a partir do qual se pode esperar que os indivíduos que sobrevivam até esse instante estejam curados. Mais especificamente, quanto mais pequeno for o valor de λ , maior o valor de t a partir da qual a função de risco é aproximadamente zero.

3.3 Estimação dos parâmetros

Caso geral

Recorde-se que, na Análise de Sobrevida, se for admitido censura não informativa e censura à direita, a função de verosimilhança para uma amostra de dimensão n pode ser expressa por

$$L = \prod_{i=1}^n f(t_i)^{\delta_i} S(t_i)^{1-\delta_i},$$

onde δ_i é uma variável indicatriz tal que $\delta_i = 1$ se t_i corresponde a um tempo de vida e $\delta_i = 0$ se t_i é uma observação censurada.

Considere-se uma amostra de dimensão n e designe-se por t_1, \dots, t_n os tempos de vida. Sem perda de generalidade, suponha-se que os primeiros m ($m < n$) tempos de vida são censurados. Sejam $\delta_1, \dots, \delta_n$ tais que

$$\delta_i = \begin{cases} 0 & \text{se } 1 \leq i \leq m \\ 1 & \text{se } m+1 \leq i \leq n \end{cases}$$

e y_1, \dots, y_n tais que

$$y_i = \begin{cases} 0 & \text{se o indivíduo é imune} \\ 1 & \text{se o indivíduo é suscetível} \end{cases}$$

Tendo em conta os valores possíveis que o par (δ_i, y_i) pode assumir, o tipo de contribuição para a verosimilhança da respetiva observação varia, como se pode observar na Tabela 1.

Tabela 1: Resumo das possibilidades para (δ_i, y_i) .

Valores possíveis para (δ_i, y_i)	Descrição da situação	Contribuição para a verosimilhança
(0, 0)	Censurado, imune	p
(0, 1)	Censurado, suscetível	$(1-p)S(t Y=1)$
(1, 1)	Observado, suscetível	$(1-p)f(t Y=1)$

No caso do modelo de cura de mistura, a verosimilhança total observada é

$$L_O = \prod_{i=1}^n [(1-p)f(t|Y=1)]^{\delta_i} [p + (1-p)S(t|Y=1)]^{1-\delta_i}.$$

Se todos os y_i 's fossem observados (em geral, não o são aqueles que correspondem a observações censuradas), pela Tabela 1 ter-se-ia a seguinte verosimilhança completa

$$L_C = \prod_{i=1}^n \left[[(1-p)f(t|Y=1)]^{y_i} \right]^{\delta_i} \left[p^{1-y_i} [(1-p)S(t|Y=1)]^{y_i} \right]^{1-\delta_i}.$$

Tendo em conta que $f(t|Y=1) = h(t|Y=1)S(t|Y=1)$ e que, dados os valores possíveis para (δ_i, y_i) tem-se $(1-\delta_i)(1-y_i) = 1 - y_i$, a verosimilhança anterior pode ser fatorizada em

$$L_C = \prod_{i=1}^n (1-p)^{y_i} p^{1-y_i} \prod_{i=1}^n h(t|Y=1)^{y_i \delta_i} S(t|Y=1)^{y_i}.$$

O facto de nem todos os y_i 's serem observados origina uma situação de dados incompletos, pelo que o algoritmo EM é um método iterativo adequado para obter as estimativas de máxima verosimilhança dos parâmetros da distribuição.

A etapa E deste algoritmo consiste em determinar o valor esperado do logaritmo da verosimilhança completa em relação à distribuição dos Y 's não observados, condicional aos valores atuais dos parâmetros e aos dados observados \mathcal{O} , onde $\mathcal{O} = \{y_i \text{ observado}, (t_i, y_i), i=1, \dots, n\}$. No entanto, como em relação às observações censuradas o logaritmo de L_C é linear em Y , para calcular o valor esperado de $\log L_C$ basta substituir, na verosimilhança completa, os valores não observados de Y pelos respectivos valores esperados, denotados por t_i . Então, tem-se:

$$\tau_i = E(Y | \mathcal{O}) = P(Y_i = 1 | T_i > t_i, \delta_i = 0) = \frac{(1-p)S(t_i | Y_i = 1)}{S(t_i)}.$$

Assim, na verosimilhança completa, cada y_i é substituído por ω_i , onde ω_i é definido na forma que se segue:

$$\omega_i = \begin{cases} 1 & \text{se } \delta_i = 1 \\ \tau_i & \text{se } \delta_i = 0 \end{cases}.$$

Na etapa E, cada observação censurada i é atribuída à subpopulação $Y=1$ com probabilidade τ_i e à subpopulação $Y=0$ com probabilidade $1 - \tau_i$.

Depois de substituir y_i por τ_i na verosimilhança completa, obtém-se a seguinte verosimilhança "esperada"

$$L_E = \prod_{i=1}^n q^{\omega_i} (1-q)^{1-\omega_i} \prod_{i=1}^n h(t|Y=1)^{\omega_i \delta_i} S(t|Y=1)^{\omega_i} = L_{E_1} L_{E_2},$$

onde $q=1-p$.

Para efetuar a etapa M, é necessário maximizar as duas componentes do logaritmo da função de verosimilhança esperada para, se possível, obter as expressões dos estimadores dos parâmetros. Relativamente ao parâmetro q , a expressão do seu estimador é dada por

$$\hat{q} = \frac{1}{n} \left[(n-m) + \sum_{i=1}^m \tau_i \right].$$

Em termos gerais, a estratégia a adotar para a implementação do algoritmo EM, resume-se a:

1. determinar os valores iniciais;
2. calcular ω_i com base nos valores atuais dos parâmetros;
3. maximizar $\log L_{E_1}$ e $\log L_{E_2}$ com base nos valores atuais de ω_i e dos parâmetros;
4. repetir 2. e 3. até à convergência.

Caso da distribuição de Chen

De acordo com a definição dos valores esperados dos valores não observados de Y , no caso de o tempo de vida dos indivíduos suscetíveis ter uma distribuição de Chen, tem-se

$$\tau_i = \frac{q \exp \left[\lambda \left(1 - \exp \left(t_i^\beta \right) \right) \right]}{1 - q + q \exp \left[\lambda \left(1 - \exp \left(t_i^\beta \right) \right) \right]}.$$

Assim, como $\omega_i = 1$ se $\delta_i = 1$ e $\omega_i = \tau_i$ se $\delta_i = 0$, a referida verosimilhança “esperada” é

$$L_E = \prod_{i=1}^n q^{\omega_i} (1-q)^{1-\omega_i} \prod_{i=1}^n \left[\lambda \beta t_i^{\beta-1} \exp \left(t_i^\beta \right) \right]^{\rho_i \delta_i} \left[\exp \left[\lambda \left(1 - \exp \left(t_i^\beta \right) \right) \right] \right]^{\omega_i} = L_{E_1} L_{E_2}$$

e os logaritmos de L_{E_1} e L_{E_2} são, respetivamente,

$$\log L_{E_1} = (n-m) \log q + m \log (1-q) + (\log q - \log (1-q)) \sum_{i=1}^m \tau_i$$

e

$$\begin{aligned} \log L_{E_2} = & \lambda \sum_{i=1}^m \tau_i \left[1 - \exp \left(t_i^\beta \right) \right] + (n-m) (\log \lambda + \log \beta) + (\beta-1) \sum_{i=m+1}^n \log t_i + \\ & + \sum_{i=m+1}^n t_i^\beta + \lambda \sum_{i=m+1}^n \left[1 - \exp \left(t_i^\beta \right) \right], \end{aligned}$$

tendo em conta que as m observações censuradas estão indexadas de 1 a m e os $n-m$ tempos de vida observados estão indexados de $m+1$ a n .

Após derivar $\log L_{E_2}$, é possível obter uma expressão explícita apenas para o estimador de λ

$$\hat{\lambda} = \frac{n-m}{\sum_{i=1}^m \tau_i \left[\exp \left(t_i^\beta \right) - 1 \right] + \sum_{i=m+1}^n \left[\exp \left(t_i^\beta \right) - 1 \right]}$$

Assim, a estimativa de β é obtida recorrendo ao método de Newton-Raphson.

Para proceder à estimação dos parâmetros, ou seja, para usar o algoritmo EM, é necessário substituir τ_i pelo respectivo valor, donde obtém-se

$$\hat{q} = \frac{1}{n} \left[(n-m) + q \sum_{i=1}^m \frac{\exp \left[\lambda \left(1 - \exp \left(t_i^\beta \right) \right) \right]}{1 - q + q \exp \left[\lambda \left(1 - \exp \left(t_i^\beta \right) \right) \right]} \right],$$

$$\hat{\lambda} = \frac{n-m}{q \sum_{i=1}^m \frac{\exp \left[\lambda \left(1 - \exp \left(t_i^\beta \right) \right) \right]}{1 - q + q \exp \left[\lambda \left(1 - \exp \left(t_i^\beta \right) \right) \right]} \left[\exp \left(t_i^\beta \right) - 1 \right] + \sum_{i=m+1}^n \left[\exp \left(t_i^\beta \right) - 1 \right]}.$$

3.4 Aplicação a dados reais

A metodologia anterior é ilustrada com a aplicação a dois conjuntos de dados reais, que podem ser encontrados na literatura: Maller e Zhou (1996, p.82) e Klein e Moeschberger (1998, p.10). Os ajustamentos foram efetuados recorrendo a algoritmos programados na linguagem R (R, 2010).

O primeiro desafio que surge diz respeito à obtenção dos valores iniciais para os parâmetros. Note-se que existem dois tipos de parâmetros: o parâmetro q e os parâmetros da distribuição de Chen, logo duas abordagens distintas. Para q , com base em Maller e Zhou (1992), considera-se a estimativa da proporção de suscetíveis $q^{(0)} = 1 - \hat{S}(t_{(r)})$, onde $t_{(r)}$ representa a maior observação não censurada e $\hat{S}(t)$ a estimativa de Kaplan-Meier da f.s. Para os parâmetros da distribuição, sugere-se duas opções:

1. efetuar um ajustamento apenas com os dados com valores até $t_{(r)}$ e as estimativas obtidas constituem os valores iniciais; este método requer a programação de um algoritmo para a distribuição de Chen;
2. determinar o valor da estimativa de Kaplan-Meier da f.s. em dois pontos afastados e resolver em ordem a λ para dois valores fixos de β ; o valor de β que origine valores de λ mais próximos é a estimativa inicial de β ; a estimativa inicial de λ é a média dos dois valores próximos.

Dados de leucemia (Maller e Zhou, 1996)

Num estudo efetuado ao longo de 5 anos, Kersey *et al.* (1987) compararam o efeito do tratamento da leucemia linfoblástica aguda através de transplante de medula óssea em 90 doentes, 46 dos quais com a sua própria medula (grupo 2) e 44 com medula de algum parente compatível (grupo 1). O tempo de vida foi definido como o tempo, em anos, desde o transplante até à recaída.

Ambos os grupos apresentam uma considerável quantidade de observações censuradas para além do maior tempo de vida observado. Como tal, optou-se por exemplificar o ajustamento apenas para o grupo 1. A partir dos valores iniciais dos parâmetros $q^{(0)}=0.73662$, $\lambda^{(0)}=0.9930455$ e $\beta^{(0)}=1$, obteve-se as estimativas que se seguem $\hat{q} = 0.7285978$, $\hat{\lambda} = 0,7611151$ e $\hat{\beta} = 0.6139747$, as quais deram origem à curva representada na Figura 2. Nesta mesma figura, encontra-se o ajustamento com o modelo de cura baseado noutras distribuições, nomeadamente a distribuição de Weibull (Maller e Zhou, 1996, p.108) e a distribuição de Burr de tipo XII (Shao e Zhou, 2004). Deste modo, é evidente que, para estes dados, o modelo de cura é adequado e que o modelo proposto (Abreu, 2005 e Abreu e Rocha, 2006) é uma boa alternativa aos modelos já existentes.

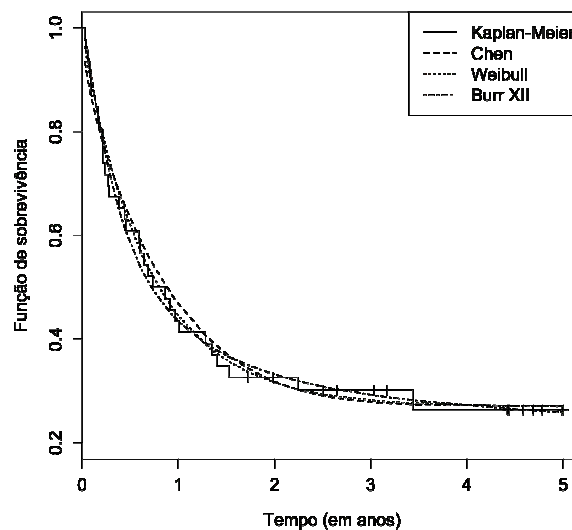


Figura 2: Dados de leucemia (Maller e Zhou, 1996, p.82)

Outros dados de leucemia (Klein e Moeschberger, 1998)

Os autores referem um estudo envolvendo 101 pacientes com leucemia mielógena aguda avançada. A todos estes pacientes foi feito transplante de medula óssea, 51 com a sua própria medula (os do grupo designado por *auto*) e 50 com medula de algum parente compatível (os do grupo designado por *allo*). Como o grupo *allo* tem uma apreciável quantidade de observações censuradas correspondentes aos valores mais elevados das observações, apresenta boas perspectivas de conter indivíduos imunes, daí que o ajustamento é realizado apenas para este grupo. De facto, o nivelamento da estimativa de Kaplan-Meier da f.s. ocorre após aproximadamente 20 meses (Figuras 1 e 3), num valor próximo de 0.5, o que é uma forte indicação da provável existência de indivíduos imunes.

Os valores iniciais dos parâmetros são $q^{(0)}=0.46786$, $\lambda^{(0)}=0.183395$ e $\beta^{(0)}=0.5$, dos quais resultaram as seguintes estimativas: $\hat{q} = 0.47060$, $\hat{\lambda} = 0.12508$, e $\hat{\beta} = 0.42100$. A curva resultante está patente na Figura 3, onde se pode notar um ajustamento quase perfeito.

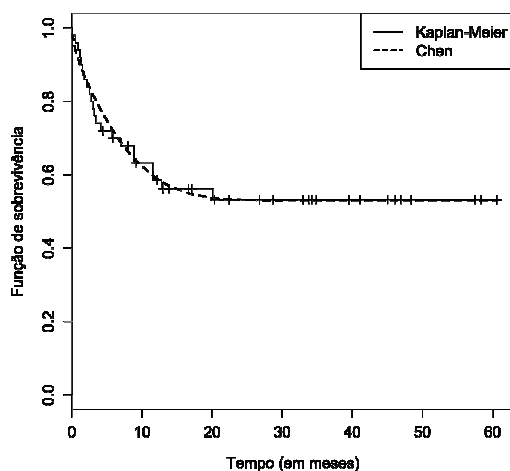


Figura 3: Dados de leucemia (Klein e Moeschberger, 1998, p.10)

4. Comentários finais

O desenvolvimento da medicina, nomeadamente no que diz respeito à cura de doenças do foro oncológico, tem originado um forte interesse na aplicação de modelos de cura, em detrimento ou como complemento dos modelos mais tradicionais. Assim, o foco passou a estar não só na estimação do tempo de vida dos indivíduos (onde se inclui o efeito das covariáveis relevantes) mas também na probabilidade de cura dos pacientes. Esta informação adicional é importante pois pode auxiliar os oncologistas na decisão da terapêutica a aplicar a cada doente; eventualmente, indivíduos com maior probabilidade de cura podem não ser sujeitos a tratamentos tão agressivos (Hunsberger *et al.*, 2009).

Importa notar que os modelos de cura já estão numa fase de desenvolvimento mais avançada do que a aqui apresentada. Por exemplo, no que diz respeito aos modelos de cura de mistura, permitem a inclusão de covariáveis quer na f.s. dos indivíduos suscetíveis, quer na probabilidade de cura.

Referências Bibliográficas

- Abreu, A.M. (2005). *Modelos de Sobrevivência para Populações com Indivíduos Imunes*. Dissertação de doutoramento. Universidade da Madeira.
- Abreu, A.M. e Rocha, C.S. (2006). *Um novo modelo de cura paramétrico*. Em *Ciência Estatística*. Editores: Castro, L.C., Martins, E.G., Rocha, C., Oliveira, M.F., Leal, M.M., Rosado, F., p. 151-162. Edições SPE, Lisboa
- Betensky, R.A. e Schoenfeld, D.A. (2001). Nonparametric estimation in a cure model with random cure times. *Biometrics*, Vol.57, p. 282-286.
- Cantor, A.B. e Shuster, J.J. (1992). Parametric versus non-parametric methods for estimating cure rates based on censored survival data. *Statistics in Medicine*, Vol.11, p. 931-937.
- Chen, Z. (2000). A new two-parameter lifetime distribution with bathtub shape or increasing failure rate function. *Statistics and Probability Letters*, Vol.49, p. 155-161.
- Hunsberger, S., Albert, P.S. e London, W.B. (2009). A finite mixture survival model to characterize risk groups of neuroblastoma. *Statistics in Medicine*, Vol.28, p. 1301-1314.
- Kersey, J.H., Weisdorf, D., Nesbit, M.E., Lebien, T.W., Woods, G.W., McGlave, P.B., Kim, T., Vallera, D.A., Goldman, A.I., Bostrom, B., Hurd, D. e Ramsay, N.K.C. (1987). Comparison of Autologous and Allogeneic Bone Marrow Transplantation for Treatment of High-Risk Refractory Acute Lymphoblastic Leukemia. *The New England Journal of Medicine*, Vol.317, p. 461-467.
- Klein, J.P. e Moeschberger, M.L. (1998). *Survival Analysis. Techniques for Censored and Truncated Data*. 2ª impressão corrigida. Springer-Verlag, New York.
- Laska, E.M. e Meisner, M.J. (1992). Nonparametric estimation and testing in a cure model. *Biometrics*, Vol.48, p. 1223-1234.
- Maller, R.A. e Zhou, S. (1992). Estimating the proportion of immunes in a censored sample. *Biometrika*, Vol.79, p. 731-739.
- Maller, R.A. e Zhou, S. (1996). *Survival Analysis with Long-Term Survivors*. John Wiley, New York.
- R Development Core Team (2010). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org/> (em 16/7/2011).
- Rocha, C.S. (1995). *Modelos com fragilidade em Análise de Sobrevivência*. Dissertação de doutoramento. Faculdade de Ciências da Universidade de Lisboa.
- Shao, Q. e Zhou, X. (2004). A new parametric model for survival data with long-term survivors. *Statistics in Medicine*, Vol.23, p. 3525-3543.
- Tamura, R.N., Faries, D.E. e Feng, J. (2000). Comparing time to onset of response in antidepressant clinical trials using the cure model and the Cramer-von Mises test. *Statistics in Medicine*, Vol.19, p. 2169-2184.
- Taylor, J.M.G. (1995). Semi-parametric estimation in failure time mixture models. *Biometrics*, Vol. 51, p. 899-907.
- Yakovlev, A.Y., Asselain, B., Bardou, V.J., Fourquet, A., Hoang, R., Rochefediere, A. e Tsodikov, A.D. (1993). *A simple stochastic model of tumor recurrence and its applications to data on premenopausal breast cancer*. Em *Biometrie et Analyse de Données Spatio-Temporelles*. Editores: Asselain, B., Boniface, M., Duby, C., Lopez, C., Masson, J.P.E Tranchefort, J., Vol.12, p. 66-82. Société Française de Biométrie, ENSA Rennes, France.

