Brazilian Journal of Probability and Statistics (2000), 14, pp. 113–122.

SIMULATION IN A FULLY PARAMETRIC PROPORTIONAL HAZARD MODEL WITH CHANGEPOINTS

Ana María Lara Porras, Julia García Leal and
Esteban Navarrete Álvarez
University of Granada
Department of Statistics and Operations Research
Campus de Fuentenueva. 18071-Granada. Spain
alara@goliat.ugr.es

Summary

A fully parametric proportional hazard survival model is studied through the simulation of the behavior of maximum-likelihood estimators. We consider the data to be generated by a Gompertz distribution with one changepoint and certain observations to be right-censored. This study has been carried out using the Mathematica program for data generation and for the numerical solution of equations, and the Statgraphic and S-Plus packages for the statistical analysis of results.

Key Words: Censored data; changepoints; covariates; Gompertz distribution; maximum likelihood; proportional hazard model.

1 Introduction

Proportional hazard models comprise the basis of the most common procedures in survival analysis. A review of the different study procedures on these models appears in Kay (1977), and Aitkin and Clayton (1980). An important improvement over these models was achieved by Noura and Read (1990), who considered that the parameters characterizing the base-line distribution may vary with time for different intervals but remain constant at each interval. The points where these changes take place are termed changepoints.

We have employed a changepoint model, considering the survival time to be determined by the Gompertz distribution.

In some practical situations if we could consider changepoints it can provide a description most specifies of the data. For example, in the medical treatment of a disease, the appearance of a new drug can affect to the time of survival. In our model, this will be reflected only in a change of the parameters that characterize to the base-line distribution. In general terms, when real changes in the patterns of risk can be recognized, changepoints may be justified physically.

In this paper we have studied by simulation the maximum-likelihood estimators for the proposed model. Due to the extreme difficulty in solving maximum-likelihood equations, it has been necessary to approach them by numerical methods.

2 The model

Let us consider a proportional hazard survival model whose survival time, t, follows Gompertz's distribution. This distribution describes a rather precise form for the length of humane life after the age of 20.

We suppose that for each individual is defined a $p \times 1$ vector $\mathbf{z} = (z_1, \cdots, z_p)^T$ of covariates which represent the characteristics which could have an influence on failure time.

The covariates are added to the distribution of failure time via the link function $\psi(\mathbf{z})$, which can be parameterized by $\psi(\mathbf{z}) = e^{\beta^T \mathbf{z}}$, where the linear predictor $\beta^T \mathbf{z}$ expresses the relative effects of the covariates $\mathbf{z}$ in terms of a vector of regression coefficients $\beta = (\beta_1, \cdots, \beta_p)^T$.

In the presence of covariates, its survivor function $S(t; \mathbf{z})$ and hazard function $\lambda(t; \mathbf{z})$ are defined by:

$$\begin{aligned} S(t;\mathbf{z}) &= \exp\left[-\alpha(e^{\rho t}-1)/\rho\right]^{e^{\beta^T \mathbf{z}}} \quad , \\ \lambda(t;\mathbf{z}) &= \alpha \exp\left[\beta^T \mathbf{z} + \rho t\right] \quad , \end{aligned} \tag{2.1}$$

where α y ρ are Gompertz's distribution parameters.

Let a partition of the time axis be given by parameter changepoints $a_1, \cdots, a_k$, with $a_0 = 0$ and $a_{k+1} = \infty$; in each interval (a_{j-1}, a_j), the distribution parameters take the values α_j and ρ_j.

Let $g(t)$ be the logarithm of the accumulative base-line hazard function, $\Lambda(t)$, given in this case by:

$$\Lambda(t) = \int_0^t \lambda_0(u)du = \int_0^t \alpha e^{\rho u} du \quad .$$

At each interval $a_{j-1} < t \leq a_j$ of the time axis, $g(t)$ becomes

$$g(t) = \ln\left[\alpha_j \left(e^{\rho_j t} - 1\right)/\rho_j\right] \quad , \quad j = 1, \ldots, k+1 \quad . \tag{2.2}$$

Imposing the continuity condition on $g(t)$ at the changepoints, the following is verified:

$$\ln\left[\frac{\alpha_j \left[e^{\rho_j a_j} - 1\right)}{\rho_j}\right] = \ln\left[\frac{\alpha_{j+1}\left(e^{\rho_{j+1} a_j} - 1\right)}{\rho_{j+1}}\right] \quad , \quad j = 1, \ldots, k \quad , \tag{2.3}$$

from which we obtain the following parameter relationship:

$$\alpha_j \quad = \frac{\alpha_1 \rho_j}{\rho_1} \prod_{p=1}^{j-1} \frac{e^{\rho_p a_p} - 1}{e^{\rho_{p+1} a_p} - 1} \quad , \quad j = 2, \ldots, k+1 \quad . \tag{2.4}$$

So for a survival time ending at j-th interval,

$$g(t) = \ln\left[\frac{\alpha_1}{\rho_1}\left(e^{\rho_j t} - 1\right)\prod_{p=1}^{j-1}\frac{e^{\rho_p a_p} - 1}{e^{\rho_{p+1} a_p} - 1}\right] \quad , \tag{2.5}$$

for i-th observation we obtain

$$g(t_i) = \sum_{j=1}^{k+1} c_{ij} \ln\left[\frac{\alpha_1}{\rho_1}\left(e^{\rho_j t_i} - 1\right)\prod_{p=1}^{j-1}\frac{e^{\rho_p a_p} - 1}{e^{\rho_{p+1} a_p} - 1}\right] \quad , \tag{2.6}$$

where c_{ij} is a variable indicator associated with the i-th observation, defined as

$$c_{ij} = \begin{cases} 1 & \text{if } a_{j-1} < t_i \leq a_j \\ 0 & \text{otherwise} \end{cases}$$

with $\quad i = 1, \ldots, N \quad$ and $\quad j = 1, \ldots, k+1 \;$. N represents the number of individuals studied.

Let $H_i = \exp\{g(t_i) + \beta^T z\}$ and

$$h_i = H_i' = g'(t_i) H_i = H_i \prod_{j=1}^{k+1}\left[\frac{\rho_j e^{\rho_j t_i}}{e^{\rho_j t_i} - 1}\right]^{c_{ij}} \quad . \tag{2.7}$$

Survivor and density functions are respectively represented by

$$\begin{aligned} S(t_i) &= \exp\left[-H_i\right] \\ f(t_i) &= h_i \exp\left[-H_i\right] \quad . \end{aligned} \tag{2.8}$$

3 Likelihood equations

Let $T_1, \ldots, T_N$ be the associated survival times of N selected individuals; these survival times may or may not be right-censored. Let $t_1, \ldots, t_N$ be the observed survival times and $\omega_1, \ldots, \omega_N$ the corresponding censoring indicators defined by

$$\omega_i = \begin{cases} 1 & \text{if the observation is not censored } (T_i = t_i) \\ 0 & \text{if it is censored } (T_i > t_i) \end{cases} \tag{3.1}$$

If the right-censored scheme is independent of the failure mechanism, it is clear that a censored value t_i contributes only the information that T_i exceeds t_i. It follows that a survival time censored in this manner contributes to the likelihood of its survivor function and an uncensored observation contributes its density function . In this manner, the likelihood function of N observations, aided by the censored indicator, is expressed as

$$l = \prod_{i=1}^{N}\left[f(t_i; \mathbf{z})\right]^{\omega_i}\left[S(t_i; \mathbf{z})\right]^{1-\omega_i} \quad , \tag{3.2}$$

and the log-likelihood as

$$\begin{aligned}
L = \quad & \sum_{i=1}^{N} \{\omega_i \ln \lambda(t_i;\mathbf{z}) + \ln S(t_i;\mathbf{z})\} = \\
& \sum_{i=1}^{N} \Bigg\{ \omega_i \ln \Bigg\{ \exp \Bigg[\sum_{s=1}^{p} \beta_s z_{is} + \sum_{j=1}^{k+1} c_{ij} \ln \Bigg[\frac{\alpha_1}{\rho_1} \left(e^{\rho_j t_i} - 1 \right) \times \\
& \prod_{p=1}^{j-1} \frac{e^{\rho_p a_p} - 1}{e^{\rho_{p+1} a_p} - 1} \Bigg] \Bigg] \prod_{j=1}^{k+1} \left[\frac{\rho_j e^{\rho_j t_i}}{e^{\rho_j t_i} - 1} \right]^{c_{ij}} \Bigg\} - \\
& \exp \Bigg\{ \sum_{s=1}^{p} \beta_s z_{is} + \sum_{j=1}^{k+1} c_{ij} \ln \Bigg[\frac{\alpha_1}{\rho_1} \left(e^{\rho_j t_i} - 1 \right) \prod_{p=1}^{j-1} \frac{e^{\rho_p a_p} - 1}{e^{\rho_{p+1} a_p} - 1} \Bigg] \Bigg\} \Bigg\} \qquad .
\end{aligned} \tag{3.3}$$

From this point, the likelihood equations associated with the model are obtained by the following three groups of expressions:

$$\begin{aligned}
\frac{\partial L}{\partial \rho_1} \quad & = \sum_{i=1}^{N} \Bigg\{ \omega_i \Bigg[c_{i1} t_i + \left(\frac{a_1 e^{\rho_1 a_1}}{e^{\rho_1 a_1} - 1} - \frac{1}{\rho_1} \right) \sum_{p=2}^{k+1} c_{ip} \Bigg] - \\
& - H_i \Bigg[c_{i1} \left(\frac{t_i e^{\rho_1 t_i}}{e^{\rho_1 t_i} - 1} - \frac{1}{\rho_1} \right) + \left(\frac{a_1 e^{\rho_1 a_1}}{e^{\rho_1 a_1} - 1} - \frac{1}{\rho_1} \right) \sum_{p=2}^{k+1} c_{ip} \Bigg] \Bigg\} = 0 \quad ,
\end{aligned} \tag{3.4}$$

$$\begin{aligned}
\frac{\partial L}{\partial \rho_j} \quad & = \sum_{i=1}^{N} \Bigg\{ \omega_i \Bigg[c_{ij} \left(\frac{1 + \rho_j t_i}{\rho_j} - \frac{a_{j-1} e^{\rho_j a_{j-1}}}{e^{\rho_j a_{j-1}} - 1} \right) + \Bigg(\frac{a_j e^{\rho_j a_j}}{e^{\rho_j a_j} - 1} - \\
& - \frac{a_{j-1} e^{\rho_j a_{j-1}}}{e^{\rho_j a_{j-1}} - 1} \Bigg) \sum_{p=j+1}^{k+1} c_{ip} \Bigg] - H_i \Bigg[c_{ij} \left(\frac{t_i e^{\rho_j t_i}}{e^{\rho_j t_i} - 1} - \frac{a_{j-1} e^{\rho_j a_{j-1}}}{e^{\rho_j a_{j-1}} - 1} \right) + \\
& + \left(\frac{a_j e^{\rho_j a_j}}{e^{\rho_j a_j} - 1} - \frac{a_{j-1} e^{\rho_j a_{j-1}}}{e^{\rho_j a_{j-1}} - 1} \right) \sum_{p=j+1}^{k+1} c_{ip} \Bigg] \Bigg\} = 0 \quad ,
\end{aligned} \tag{3.5}$$

for $j = 2, \ldots, k+1$, and

$$\frac{\partial L}{\partial \beta_s} \quad = \sum_{i=1}^{N} (\omega_i - H_i) z_{is} = 0 \quad , \qquad\qquad s = 0, \ldots, p \quad , \tag{3.6}$$

the term $\ln \alpha_1$ is identified with the component β_0 of the parameter vector β which requires the addition of a component $z_{i0} = 1$ to the covariates vector.

4 Simulation scheme

In order to determine the behavior of the estimators resulting from equations (3.4), (3.5), and (3.6), we have performed a simulation study in which we considered the following simplifications:

- A single changepoint.
- The existence of two groups of individuals, taking the same number of observations in each group.

We have also performed a comparative study varying the following factors:

- Changepoint position (percentiles 40 and 60).
- Sample size (50, 80, and 200).

Throughout the study we have considered that the probability of a censored individual is 0.2. To each individual, a random number is assigned generated by a uniform distribution; if said number is less then 0.2, the individual is censored and his observed survival time is randomly reduced by the following expression:

$$F(t_i) = \gamma_i * F(T_i) \quad \text{with} \quad \gamma_i \in U[0,1].$$

In other words, the reduction of survival time of censored individuals is modeled by an uniform distribution.

The determination of the estimators via the Mathematica program was carried out by numerically solving the maximum-likelihood equations obtained from the explicit expression of the corresponding derivatives.

5 Simulation results

Taking into account only the changepoint variation we distinguish two situations, I and II. In both cases, we assign parameters $\alpha_1 = 0.01$ and $\rho_1 = 1$ (an arbitrary choice of values) to the distribution before the changepoint. The selection of ρ_2 was arbitrary. The value for α_2 was obtained from the recurring ratio (2.4). We have supposed that $\beta_1 = 0$, therefore the difference between the two groups is produced by β_2. From these assumptions each situation is identified by the parameters summarized in Table 1.

Table 1

Common characteristics		
$n_1 = n_2$	$\beta_1 = 0$	$\alpha_1 = 0.01$ $\rho_1 = 1$
% Censored = 20	$\beta_2 = 0.59$	$\rho_2 = 0.5$
Parameters	Situation I	Situation II
Changepoint	3.9 (P_{40})	4.5 (P_{60})
α_2	0.04	0.0524

The simulation is based on the study of 100 independent samples of each sample size.

A summary of the results obtained in each of the two situations for sample sizes of 50, 80 and 200 is given below in Tables 2 and 3.

Table 2

Estimator behavior. Situation I.						
	$n = 50$		$n = 80$		$n = 200$	
Param.	Bias	St.Error	Bias	St.Error	Bias	St.Error
$\hat{\beta}_2$	-0.040853	0.027499	-0.005165	0.001881	-0.00097	0.000305
$\hat{\alpha_1}$	0.000577	0.000518	-0.000484	0.000123	-0.000212	0.000104
$\hat{\alpha_2}$	0.007899	0.001266	0.007763	0.000937	-0.000577	0.001843
$\hat{\rho}_1$	-0.036411	0.011655	0.008639	0.004513	0.000383	0.000131
$\hat{\rho}_2$	-0.024669	0.002822	-0.002969	0.000863	-0.000291	0.000071

Table 3

Estimator behavior. Situation II.						
	$n = 50$		$n = 80$		$n = 200$	
Param.	Bias	St.Error	Bias	St.Error	Bias	St.Error
$\hat{\beta}_2$	-0.008321	0.007397	-0.004816	0.000878	-0.000126	0.000055
$\hat{\alpha_1}$	0.001741	0.000472	0.001256	0.000174	0.000625	0.00013
$\hat{\alpha_2}$	0.009771	0.002074	-0.000789	0.000439	-0.000079	0.000422
$\hat{\rho}_1$	0.006481	0.001973	0.005131	0.00608	-0.000237	0.000172
$\hat{\rho}_2$	-0.034685	0.003994	-0.003525	-0.000338	-0.000239	0.000198

Comparing the results obtained with the sample sizes 50, 80 and 200, we note that each estimator bias decreases when the sample size increases. The standard error also tends to zero and the estimator approaches the true value of the parameter.

For 200-sized samples, it can be statistically accepted in Situation I and Situation II that estimators $\hat{\beta}_2$, $\hat{\alpha_1}$, $\hat{\alpha_2}$, $\hat{\rho}_1$ and $\hat{\rho}_2$ are approximately unbiased.

In addition, the correlation matrices of the samples between the estimators in each of the situations and for a sample of 200 are shown in Tables 4 and 5.

Table 4

Sample correlations of estimators. Situation I					
	$\hat{\beta}_2$	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\rho}_1$	$\hat{\rho}_2$
$\hat{\beta}_2$	1.0000	- 0.0365	0.0015	0.0282	-0.0209
	0.0000	0.6076	0.9833	0.6923	0.7692
$\hat{\alpha}_1$	-0.0365	1.0000	0.0757	0.0292	0.1280
	0.6076	0.0000	0.2869	0.6813	0.0709
$\hat{\alpha}_2$	0.0015	0.0757	1.0000	0.0863	-0.0848
	0.9833	0.2869	0.0000	0.2246	0.2324
$\hat{\rho}_1$	0.0282	0.0292	0.0863	1.0000	0.0183
	0.6923	0.6813	0.2246	0.0000	0.7974
$\hat{\rho}_2$	-0.0209	0.1280	-0.0848	0.0183	1.0000
	0.7692	0.0709	0.2324	0.7974	0.0000

Table 5

Sample correlations of estimators. Situation II					
	$\hat{\beta}_2$	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\rho}_1$	$\hat{\rho}_2$
$\hat{\beta}_2$	1.0000	-0.0100	0.0036	0.0971	0.0877
	0.0000	0.8877	0.9596	0.1712	0.2170
$\hat{\alpha}_1$	-0.0100	1.0000	0.0699	-0.1307	0.0976
	0.8877	0.0000	0.3255	0.0651	0.1694
$\hat{\alpha}_2$	0.0036	0.0699	1.0000	0.0006	-0.0911
	0.9596	0.3255	0.0000	0.9934	0.1996
$\hat{\rho}_1$	0.0971	-0.1307	0.0006	1.0000	-0.0189
	0.1712	0.0651	0.9934	0.0000	0.7900
$\hat{\rho}_2$	0.0877	0.0976	-0.0911	-0.0189	1.0000
	0.2170	0.1694	0.1996	0.7900	0.0000

Two numbers appear in each cell of the matrix: the correlation coefficient estimate for the two variables represented by the cell and the significance level of the correlation.

In general terms, we can accept that the estimator of the covariate effect does not depend on the parameter estimators of the baseline distribution and that the estimators of the base distribution do not depend upon each other either. Said dependency is enhanced as the sample size increases.

On the other hand, the independence between the base distribution parameters is less significant due to the continuity condition demanded to the accumulative hazard function at the changepoints, given by the expression (2.3).

Finally, we show a graphic and analytical comparison between the model with a changepoint and without.

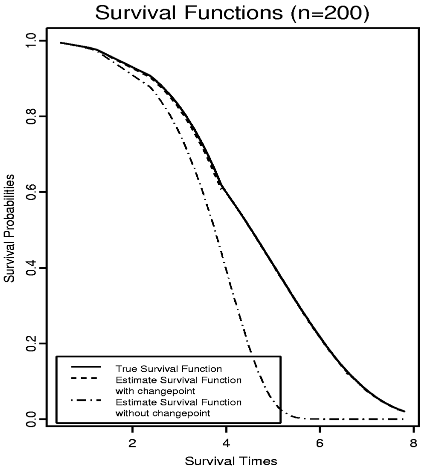

Fig. 1a: Situation I. Group 1

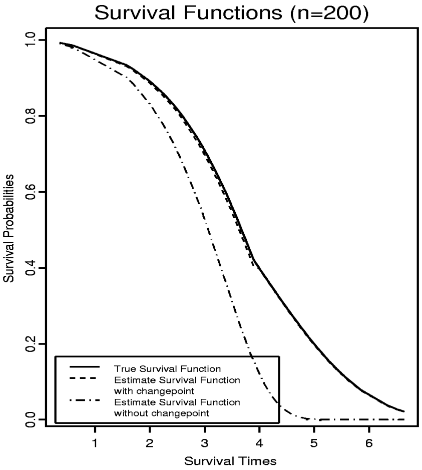

Fig. 1b: Situation I. Group 2

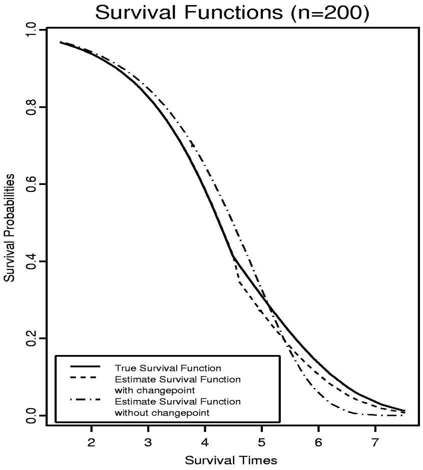

Fig. 2a: Situation II. Group 1

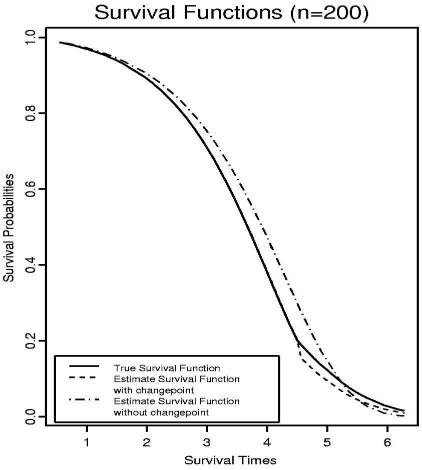

Fig. 2b: Situation II. Group 2

The figures show that the estimate survival functions with a changepoint are closer to the true survival functions than the estimates without changepoint. Also we note that the changepoints in Situation I are more important than in Situation II since the changepoints provide less information to the survival function when the changepoints occur at the end of the survival time.

The analytical, the comparison between the fit obtained by the model with and without changepoint was done by deviance analysis. For example, for data shown in Figure 1 (a and b), the two-stage Gompertz model with $a_1 = 3.9$ does significantly better than the simple Gompertz model, reducing the deviance by 96.5283, ($\chi^2 = 417.6270 - 321.0987$) on 1 DF, $p < 0.001$. Comparisons of deviance confirm that the two-stage model is also superior in Situation II.

Table 6 displays the deviances for two situations of 10 samples each. Analysis of the table shows that, in general, the model with changepoint explains the data better than the model without changepoint.

Table 6

	Deviance			
	Situation I		Situation II	
Sample	Without Chp.	With Chp.	Without Chp.	With Chp.
1	417.6270	321.0987	275.2760	230.7480
2	385.4060	302.2681	261.8940	229.4720
3	398.9920	328.9870	287.9490	260.9070
4	478.7950	329.0701	270.6319	257.79001
5	379.7082	323.1626	321.1324	279.8066
6	497.2571	321.3911	338.4481	321.9874
7	458.5673	327.0950	310.6986	294.0866
8	487.9750	389.6469	325.2892	271.3752
9	398.6766	314.0389	315.0724	291.8536
10	461.3570	350.2241	287.3906	251.5675

(Received January, 1998. Revised July, 2000.)

References

Aitkin, M. and Clayton, D. (1980). The fitting of exponential, Weibull and extreme value distributions to complex censored survival data using GLIM. *Applied Statistics*, **29**, 156-163.

Aitkin, M., Andersen, D., Francis, B. and Hinde, J. (1989). *Statistical Modelling in GLIM.* Oxford: University Press.

Cox, D.R. (1972). Regression models and life-tables (with discussion). *Journal of the Royal Statistical Society* B, **34**, 187-202.

Cox, D.R. and Oakes, D. (1984). *Analysis of Survival Data.* London: Chapman and Hall.

Kalbfleisch, J.D. and Prentice, R.L. (1980). *The Statistical Analysis of Failure Time Data.* New York: John Wiley and Sons.

Kay, R. (1977). Proportional hazard regression models and the analysis of censored survival data. *Applied Statistics*, **26**, 227-237.

Lagakos, S.W. (1981). The graphical evaluation of explanatory variables in proportional hazard regression models. *Biometrika*, **68**, 97-98.

Lara, A.M. (1995). *Aportaciones a modelos de supervivencia: distribuciones base con puntos de cambio y covariables dependientes del tiempo.* Unpublished P.H.D. dissertation. Spain: University of Granada.

McCullagh, P. and Nelder, J.A. (1989). *Generalized Linear Models.* London: Chapman and Hall.

Noura, A.A. and Read, K.L.Q (1990). Proportional hazard changepoints models in survival analysis. *Applied Statistics*, **39**, 241-253.

Prentice, R.L. (1973). Exponential survivals with censoring and explanatory variables. *Biometrika*, **60**, 279-288.