Variate Generation for Nonhomogeneous Poisson Processes with Time Dependent Covariates

Li-Hsing Shih

Lawrence Leemis

Variate Generation for a Nonhomogeneous Poisson Process with Time Dependent Covariates

LI-HSING SHIH

Department of Mineral and Petroleum Engineering
Cheng Kung University
Tainan, Taiwan

and

LAWRENCE LEEMIS

School of Industrial Engineering
The University of Oklahoma
202 West Boyd, Room 124
Norman, OK 73019

April, 1992

Algorithms are developed for generating a sequence of event times from a nonhomogeneous Poisson process that is influenced by the values of covariates that vary with time. Closed form expressions for random variate generation are shown for several baseline intensity and link functions. Two specific models linking the baseline process to the general model are considered: the accelerated time model and the proportional intensity model. In the accelerated time model, the cumulative intensity function of a nonhomogeneous Poisson process under covariate effects is $\Lambda(t;\mathbf{z}(t)) = \Lambda_0(\int_0^t \psi(\mathbf{z}(u))\,du)$, where $\mathbf{z}$ is a covariate vector, $\Lambda_0(t)$ is the baseline cumulative intensity function and $\psi(\mathbf{z})$ is the link function. In the proportional intensity model, the cumulative intensity function of a nonhomogeneous Poisson process under covariate effects is $\Lambda(t;\mathbf{z}(t)) = \int_0^t \psi(\mathbf{z}(u))\lambda_0(u)\,du$, where $\lambda_0(t)$ is the baseline intensity function.

KEY WORDS: Covariates; Nonhomogeneous Poisson processes; Simulation; Variate generation.

1. INTRODUCTION

Event times from a nonhomogeneous Poisson process with constant covariates can be generated by modifying existing algorithms. For the proportional intensity model, the cumulative intensity function under covariate effects is $\Lambda(t;\, \mathbf{z}) = \psi(\mathbf{z})\Lambda_0(t)$, where the baseline cumulative intensity function $\Lambda_0(t)$ is known, the link function $\psi(\mathbf{z})$ does not vary with time and $\mathbf{z}$ is a $q \times 1$ vector of covariates. A common link function is the log linear form $\psi(\mathbf{z}) = e^{\beta'\mathbf{z}}$, where β is a $q \times 1$ vector of regression coefficients. Generating events from the process under covariate effects is straightforward (since the cumulative intensity is multiplied by a constant $\psi(\mathbf{z})$) when an appropriate algorithm exists for generating from the baseline process.

The cumulative intensity function for the accelerated time model under covariate effects is $\Lambda(t;\, \mathbf{z}) = \Lambda_0(t\psi(\mathbf{z}))$. Event times under covariate effects can be obtained by dividing the event times that are generated for the baseline distribution by the link function since $\psi(\mathbf{z})$ does not vary with time.

Generating event times for NHPPs with time dependent covariates is more complicated than the constant covariate case. Section 2 contains a literature review on variate generation for NHPPs, survival analysis with time dependent covariates and variate generation for NHPPs with covariate effects. In Section 3, we discuss the proportional intensity and accelerated time models with time dependent covariates. Section 4 presents event time generation algorithms with time dependent covariates.

2. LITERATURE REVIEW

The literature for nonrepairable systems (i.e., survival analysis) and repairable systems (i.e., point process models) is discussed separately in the two subsections that follow. The discussion primarily concerns modeling and variate generation, and the only point process model considered is the NHPP.

2.1 Nonrepairable systems

Several authors have included time dependent covariates in survival analysis. Prentice and Kalbfleisch (1979) discussed estimation problems associated with the proportional hazards, accelerated life and competing risks models with covariates. They also consider covariates that vary with time. Dale (1985) used the proportional hazards model in the presence of time dependent covariates to model the failure times of motorettes under various temperatures. Kalbfleisch and McIntosh (1977) compared the efficiency of the partial likelihood method and Weibull analysis for the Cox proportional hazards model with time dependent covariates. Petersen (1986a, 1986b) proposed an algorithm for estimating parameters by maximum likelihood in a large variety of parametric survival models by using the Gauss-Newton method. The approach allowed for a flexible treatment of time dependent covariates.

Hoffmann (1985) considered the Weibull and piecewise constant hazard functions as baseline distributions for the proportional hazards model with time-varying covariates. He gave techniques for generating random variates from these models. Smith (1987) included cost factors in a Monte Carlo evaluation of a system of components. Leemis, Shih and Reynertson (1990) discussed random variate generation for proportional hazards and accelerated life models with time dependent covariates. For the accelerated life model, a random variate t can be generated by two consecutive inversions

$$t \leftarrow \Psi^{-1}(H_0^{-1}(-\log(1-u)))$$

where $\Psi(t) = \int_0^t \psi(u)du$ is the cumulative link function, H_0 is the baseline cumulative hazard function and u is uniformly distributed between 0 and 1. For the proportional hazards model, a random variate can be obtained by inverting the cumulative hazard function

$$H(t; \mathbf{z}(t)) = \int_0^t \psi(\mathbf{z}(u))h_0(u)du \qquad\qquad t \geq 0.$$

A closed form equation for variate generation requires inversion of $H(t; \mathbf{z}(t))$.

2.2 Repairable systems

Many simulation textbook authors (e.g., Fishman (1978), Lavenberg (1983), Law and Kelton (1991), Devroye (1986) and Ross (1990)) have suggested the use of NHPPs for modeling systems with time-varying arrival rates. Several studies used parametric intensity functions to simulate NHPPs. Lewis and Shedler (1976) proposed a method for simulating a nonhomogeneous Poisson process with log linear intensity function, $\lambda(t) = \exp(\alpha_0 + \alpha_1 t)$. Lewis and Shedler (1979a) proposed a method for simulating a nonhomogeneous Poisson process with intensity function which is a degree-two exponential polynomial, where $\lambda(t) = \exp(\alpha_0 + \alpha_1 t + \alpha_2 t^2)$. Lee, Wilson and Crawford (1991) used an exponential-trigonometric intensity function to model and simulate a cyclic storm-arrival process.

Nonparametric intensity functions are also popular for simulating NHPPs. Kaminsky and Rumpf (1977) discussed three approximate methods that are used to generate arrivals for a nonhomogeneous Poisson process and compared them to an exact method. Lewis and Shedler (1979b) proposed a general method for simulating a nonhomogeneous Poisson process by thinning. Thinning involves determining a majorizing intensity function $\lambda^*(t) \geq \lambda(t)$. The algorithm yields a series of event times from $\lambda(t)$ that are a "thinned" series of event times from $\lambda^*(t)$. Leemis (1991) proposed a piecewise linear estimator for the cumulative intensity function of an NHPP from one or more realizations. Inversion was used to generate event times for the NHPP. Other articles on the variate generation of NHPPs include Fishman and Kao (1977), Ogata (1981) and Klein and Roberts (1984).

Recently, several authors have included covariates in NHPP models (e.g., Prentice, Williams and Peterson (1981), Anderson and Gill (1982), Karr (1986) and Lawless (1987)). Two models, the proportional intensity and accelerated time models, which are analogous to proportional hazards and accelerated life models used in survival analysis are used to incorporate the covariate effects in NHPP models. Allison (1984) discussed event history analysis with time dependent covariates with applications in the social sciences. Leemis (1987) used accelerated life and proportional hazards models to incorporate the covariate effects and gave variate generation algorithms in both the renewal and nonhomogeneous Poisson process cases.

To summarize the variate generation techniques, Table 1 shows a taxonomy of variate generation for constant covariates. It includes the formulas for generating event times with constant covariates given in Leemis (1987).

	Renewal	NHPP
Accelerated life	$t \leftarrow a + \dfrac{H_0^{-1}(-\log(u))}{\psi(\mathbf{z})}$	$t \leftarrow \dfrac{\Lambda_0^{-1}(\Lambda_0(a\psi(\mathbf{z})) - \log(u))}{\psi(\mathbf{z})}$
Proportional hazards	$t \leftarrow a + H_0^{-1}(\dfrac{-\log(u)}{\psi(\mathbf{z})})$	$t \leftarrow \Lambda_0^{-1}(\Lambda_0(a) - \dfrac{\log(u)}{\psi(\mathbf{z})})$

Table 1. Formulas for generating event times with constant covariates.

3. PROPORTIONAL INTENSITY AND ACCELERATED TIME MODELS

The section defines and illustrates the proportional intensity and accelerated time models in the following two subsections.

3.1 Proportional Intensity Model

The definition of the proportional intensity model with constant covariates, $\lambda(t;\, \mathbf{z}) = \psi(\mathbf{z})\lambda_0(t)$, can be generalized for the model associated with time dependent covariates. With time dependent covariates, the link function is denoted by $\psi(\mathbf{z}(t))$ which is a multiplier of the baseline intensity function. The proportional intensity model with time dependent covariates can be defined by

$$\lambda(t;\, \mathbf{z}(t)) = \psi(\mathbf{z}(t))\lambda_0(t)$$

where $\lambda_0(t)$ is a baseline intensity function.

3.2 Accelerated Time Model

The essence of the accelerated time model in the case of covariates that do not vary with time is that "time" is contracted or expanded relative to that at $\mathbf{z}=\mathbf{0}$ (baseline case). The definition of the accelerated life model given by Cox and Oakes (1984) can also be generalized to the accelerated time model with time dependent covariates

$$dt^{\mathbf{z}} \,/\, dt^{\mathbf{0}} = 1 \,/\, \psi(\mathbf{z}(t^{\mathbf{z}}))$$

where time $t^{\mathbf{z}}$ is time for a system under $\mathbf{z}(t)$ and the corresponding "time" $t^{\mathbf{0}}$ is for the system under $\mathbf{z}(t)=\mathbf{0}$. If we integrate both sides of the expression $dt^{\mathbf{0}} = \psi(\mathbf{z}(t^{\mathbf{z}}))\, dt^{\mathbf{z}}$ from zero to the corresponding event times, we have

$$\int_0^{t^{\mathbf{0}}} du^{\mathbf{0}} = \int_0^{t^{\mathbf{z}}} \psi(\mathbf{z}(u^{\mathbf{z}}))\, du^{\mathbf{z}}$$

which yields

$$t^{\mathbf{0}} = \int_0^{t^{\mathbf{z}}} \psi(\mathbf{z}(u^{\mathbf{z}}))\, du^{\mathbf{z}}$$

where $t^{\mathbf{0}}$ is the event time under baseline conditions corresponding to $t^{\mathbf{z}}$, the event time under covariate effects. Notice that in survival analysis, this implies that $H_0(t^{\mathbf{0}}) = H(t^{\mathbf{z}})$, where H is the cumulative hazard function. This means that the cumulative hazard function at $t^{\mathbf{0}}$ of a system under baseline conditions is equal to the cumulative hazard function at $t^{\mathbf{z}}$ of a system under treatment conditions $\mathbf{z}$. In NHPP models, this is generalized to $\Lambda_0(t_i^{\mathbf{0}}) = \Lambda(t_i^{\mathbf{z}})$, $i = 1, 2, \ldots$. Here $t_i^{\mathbf{0}}$ denotes the i^{th} event time under baseline conditions and $t_i^{\mathbf{z}}$ denotes the corresponding i^{th} event time under covariate effects. A definition of accelerated time models given by Lawless (1987, page 815) for the constant covariate case is $\Lambda(t^{\mathbf{z}})=\Lambda_0(\psi(\mathbf{z})t^{\mathbf{z}})$. This definition is a special case of the above definition since

$$\Lambda(t_i^{\mathbf{z}}) = \Lambda_0(t_i^{\mathbf{0}}) = \Lambda_0(\int_0^{t_i^{\mathbf{z}}} \psi(\mathbf{z}(u^{\mathbf{z}}))du^{\mathbf{z}}) = \Lambda_0(\psi(\mathbf{z})t_i^{\mathbf{z}})$$

for $i = 1, 2, \ldots$. The interpretation is that the expected cumulative number of events at $t^{\mathbf{z}}$ under covariate effects is equal to the expected cumulative number of events at $t^{\mathbf{0}}$ under the baseline conditions. With this relationship, we can determine the equivalent event time under covariate effects by knowing the event time under baseline conditions when the link function and the time dependent covariates are known. The intensity function under covariate effects can also be obtained by using this expression and a given baseline intensity. Three examples of step, linear and power forms of a single time dependent covariate are used to illustrate the relationship between t^z and t^0. The log-linear link function $\psi(z) = e^{\beta z}$ is assumed, where β is the regression coefficient associated with the single covariate z. For simplicity, we drop the subscript "i" and use t^0 and t^z to denote the corresponding event times under baseline and covariate conditions respectively.

Example 1 (step covariate function)

First, we consider the binary step covariate case and indicate a biomedical application. Then a general step covariate function is considered with covariate values c_1 and c_2.

(a) Let the covariate be a step function with a jump at $t^z = w$

$$z(t^z) = \begin{cases} 0 & 0 \le t^z \le w \\ 1 & t^z > w. \end{cases}$$

This type of time dependent covariate might appear in a biomedical application where one level of covariate (e.g., no drug is taken by the patient) is applied up to time w and another is applied after time w (e.g., drug is taken). The event time t^0 under baseline conditions and t^z under treatment conditions might denote the times when a patient's blood pressure falls below a particular threshold. The corresponding t^0 is obtained by using $t^0 = \int_0^{t^z} \psi(z(u^z))\, du^z$,

$$t^0 = \begin{cases} t^z & 0 \le t^z \le w \\ w + e^{\beta}(t^z - w) & t^z > w. \end{cases}$$

The "baseline time", t^0, is identical to time under the influence of the covariate, t^z, up to time w. After time w, three special cases exist depending on the value of β, the regression coefficient. If $\beta = 0$, $t^0 = t^z$ for all values of t^z, which means that the change in the value of the covariate at time w has no influence on the event times. If $\beta > 0$, then t^0 is greater than t^z for $t^z > w$, indicating that the treatment accelerates the event times. Figure 1a shows that the link function $e^{0.5\, z(t)}$ is also a step function. The relationship between t^0 and t^z in this case is shown in Figure 1b for $w = 2.5$ and $\beta = 0.5$. Note that the slope of the function plotted in Figure 1b changes from 1 between $t^z = 0$ to $t^z = 2.5$ to $e^{0.5} = 1.6487$ after $t^z = 2.5$. If $\beta < 0$, then t^0 is less than t^z for $t^z > w$, indicating that the treatment decelerates the event times.

(b) Let the covariate be a step function with a jump at $t^z = w$ and

$$z(t^z) = \begin{cases} c_1 & 0 \le t^z \le w \\ c_2 & t^z > w. \end{cases}$$

This type of time dependent covariate might appear in an application where one level of covariate (e.g., the turning speed for a drill bit is 2400 RPM) is applied up to time w and another is applied after time w (e.g., the turning speed increases from 2400 RPM to 3600 RPM at time w). The corresponding t^0 is obtained by using $t^0 = \int_0^{t^z} \psi(z(u^z))\, du^z$.

$$t^0 = \begin{cases} e^{\beta c_1} t^z & 0 \le t^z \le w \\ w e^{\beta c_1} + e^{\beta c_2}(t^z - w) & t^z > w. \end{cases}$$

The "baseline time", t^0, is proportional to time under the influence of the covariate, t^z, up to time w, and the multiplicative factor is $e^{\beta c_1}$. After time w they are still proportional, however, the multiplicative factor becomes $e^{\beta c_2}$.

Example 2 (piecewise linear covariate function)

Let the covariate be a constant (equal to one) for $t^z \le w$. For $t^z \ge w$, the covariate is a linear function.

$$z(t^z) = \begin{cases} 1 & 0 \le t^z \le w \\ t^z & t^z > w. \end{cases}$$

The corresponding t^0 is obtained by using $t^0 = \int_0^{t^z} \psi(z(u^z))\, du^z$,

$$t^0 = \begin{cases} t^z e^{\beta} & 0 \le t^z \le w \\ w e^{\beta} + \dfrac{1}{\beta}(e^{\beta t^z} - e^{\beta w}) & t^z > w. \end{cases}$$

Example 3 (power covariate function)

Let the covariate be a power function, i.e., $z(t^z) = (t^z)^m$ for all $t^z \ge 0$. The general expression for t^0 is

$$t^0 = \int_0^{t^z} \psi(z(u^z))du^z = \int_0^{t^z} e^{\beta(u^z)^m}\, du^z \qquad t^z \geq 0.$$

When $m = 1$, the relationship can be expressed in closed form

$$t^0 = \frac{1}{\beta}(e^{\beta(t^z)} - 1).$$

4. GENERATION OF EVENT TIMES

To determine the appropriate method for generating event times from an NHPP with time dependent covariates, assume that the last event in an NHPP has occurred at time a. The random variable T denotes the next event time. The cumulative intensity function for the time of the next event given that the last event has occurred at time a is

$$\Lambda_{T\,|\,T>a}(t;\, \mathbf{z}(t)) = \Lambda(t;\, \mathbf{z}(t)) - \Lambda(a;\, \mathbf{z}(t)).$$

This expression allows a modeler to calculate the expected number of events after time a. A result that is useful in variate generation is given by Cinlar (1975).

> Let $N(t)$ be a counting process and let $\Lambda(t)$ be a nondecreasing function of time, t. Then $T_1, T_2, \ldots$ are the event times in a nonhomogeneous Poisson process with $E(N(t)) = \Lambda(t)$ if and only if $\Lambda(T_1), \Lambda(T_2), \ldots$ are the event times in a homogeneous Poisson process with rate 1.

According to this result, $\Lambda_{T\,|\,T>a}(t; \mathbf{z}(t))$ has a unit exponential distribution, where T is the next event time and $T - a$ is the time to the next event. If we can find expressions for $\Lambda(t;\, \mathbf{z}(t))$ for the proportional intensity and accelerated time models, the next event time can be generated by using

$$t \leftarrow \Lambda^{-1}(\Lambda(a; \mathbf{z}(t)) - \log(1 - u))$$

where u is a random number uniformly distributed on $[0, 1)$. When $\Lambda^{-1}(t)$ and $\Lambda(t)$ are closed form, the next event time can be generated by a closed form expression. A series of event times of an NHPP can be generated by using the expression repeatedly. Note that the basic form of variate generation for lifetime or the first event time of an NHPP, $t \leftarrow H^{-1}(-\log(1-u))$, is a special case of this expression since $\Lambda(a) = \Lambda(0) = 0$ and the cumulative hazard function $H(t)$ is equivalent to the cumulative intensity function $\Lambda(t)$.

For the accelerated time model, the cumulative intensity function is

$$\Lambda(t; z(t)) = \Lambda_0(\Psi(t)) = \int_0^{\Psi(t)} \lambda_0(u)\, du$$

where $\Psi(t) = \int_0^t \psi(z(u))\, du$ and λ_0 is the baseline intensity function. For the proportional intensity model, the cumulative intensity function is

$$\Lambda(t; z(t)) = \int_0^t \psi(z(u))\, \lambda_0(u) du$$

where ψ is the link function.

Table 2 presents variate generation algorithms for the time dependent covariate cases. The first column contains the algorithms given in Leemis, Shih and Reynertson (1990) for the accelerated time (AT) and proportional intensity (PI) models when generation of the first event time only or a renewal process is of interest. The second column contains the corresponding variate generation algorithms for NHPPs.

In the proportional intensity and accelerated time models, a link function is a multiplication factor of the intensity and time. It is interpreted as the proportion of contraction or expansion of the intensity and time for a system under covariate effects.

	Renewal	NHPP
AT	$t \leftarrow a + \Psi^{-1}(H_0^{-1}(-\log(1-u)))$ $\Psi(t) = \int_0^t \psi(u)du$	$t \leftarrow \Lambda^{-1}(\Lambda(a;\mathbf{z}(t)) - \log(1-u))$ $\Lambda(t;\mathbf{z}(t)) = \int_0^{\Psi(t)} \lambda_0(u)du$
PI	$t \leftarrow a + H^{-1}(-\log(1-u))$ $H(t;\mathbf{z}(t)) = \int_0^t \psi(\mathbf{z}(u))h_0(u)du$	$t \leftarrow \Lambda^{-1}(\Lambda(a;\mathbf{z}(t)) - \log(1-u))$ $\Lambda(t;\mathbf{z}(t)) = \int_0^t \psi(\mathbf{z}(u))\lambda_0(u)du$

Table 2. Formulas for generating event times with time dependent covariates.

There are two approaches to study variate generation with time dependent covariates. The first approach is to assume that the link function itself is a function of time, $\psi(\mathbf{z}(t)) = \psi(t)$. The second approach is to assume a time dependent covariate function $\mathbf{z}(t)$, then construct the link function based on the commonly used log linear form $e^{\beta'\mathbf{z}(t)}$. The advantage of using the link function as a function of time is that the inversion of cumulative intensity function is often more mathematically tractable. On the other hand, the advantage of defining the covariates as a function of time is that the physical measurement of the covariates can be directly applied to the models. The first approach is used in this section.

Two types of time dependent link functions are used to illustrate the event times generation algorithms for a single time dependent covariate. Four types of baseline cumulative intensity functions are assumed associated with the covariate effects in the models. Note that we use *baseline distribution* for the condition with $\mathbf{z}(t) = \mathbf{0}$ in the survival models and *baseline process* for the condition with $\mathbf{z}(t) = \mathbf{0}$ in the NHPP models. Table 3 shows the types of the link functions $\psi(t)$, the corresponding covariate functions $\mathbf{z}(t)$ and the researchers that used them.

link function $\psi(t)$	covariate $z(t)$	used by
step	step	Dale (1985), Peterson (1986a)
exponential	linear	Prentice and Kalbfleisch (1979), Peterson (1986a)
power	logarithm	Prentice and Kalbfleisch (1979), Kalbfleisch and McIntosh (1977)

Table 3. Relationships between the time dependent link function and the covariate function.

In the following discussion, a single time dependent covariate is considered. When the form of the link function is known, the corresponding covariate function is obtained by $\psi(t) = e^{\beta z(t)}$. For example, when a step link function is defined by

$$\psi(z(t)) = \begin{cases} c_1 & 0 \le t < b \\ c_2 & t \ge b \end{cases}$$

where b, c_1 and c_2 are constants. If β is the regression coefficient, then the corresponding covariate function is

$$z(t) = \begin{cases} \log c_1 / \beta & 0 \le t < b \\ \log c_2 / \beta & t \ge b \end{cases}$$

which is also a step function. When the link function has exponential form, $\psi(t) = e^{\beta t}$, the corresponding covariate is a linear function, $z(t) = t$, for $t \ge 0$. When the link function is a power function, $\psi(t) = t^k$, the corresponding covariate is a logarithmic function, $z(t) = \frac{k}{\beta} \log t$, for $t \geq 0$.

In the next two subsections, step and exponential link functions are used with four baseline processes to illustrate the algorithms for the accelerated time and proportional intensity models. The baseline processes are the homogeneous Poisson process, the power law process and the processes with log logistic and exponential power intensity forms. A homogeneous Poisson process is defined by the cumulative intensity function

$$\Lambda_0(t) = \lambda t \qquad\qquad t \geq 0$$

where λ is the rate of occurrence of failures. For the power law process, the parameterization of the cumulative intensity function is

$$\Lambda_0(t) = \nu\, t^{\delta} \qquad\qquad t \geq 0$$

where ν is a scale parameter and δ is a shape parameter. For the log logistic process, the cumulative intensity and the intensity functions are

$$\Lambda_0(t) = \log(1 + (\rho t)^{\kappa}) \qquad\qquad \lambda_0(t) = \frac{\rho\kappa(\rho t)^{\kappa - 1}}{1 + (\rho t)^{\kappa}} \qquad\qquad t \geq 0$$

where ρ is a scale parameter and κ is a shape parameter. For the exponential power process, the cumulative intensity and the intensity functions are

$$\Lambda_0(t) = e^{(t/\tau)^{\gamma}} - 1 \qquad\qquad \lambda_0(t) = \frac{\gamma}{\tau}\left(\frac{t}{\tau}\right)^{\gamma - 1} e^{(t/\tau)^{\gamma}} \qquad\qquad t \geq 0$$

where τ is a scale parameter and γ is a shape parameter. All parameters in these four baseline processes are positive.

4.1 Accelerated Time Model

For the accelerated time model, the cumulative intensity function with covariates is

$$\Lambda(t;z(t)) = \Lambda_0(\Psi(t)) = \int_0^{\Psi(t)} \lambda_0(u)\,du$$

where $\Psi(t) = \int_0^t \psi(\mathbf{z}(u))\,du$, ψ is the link function and λ_0 is the baseline intensity function. The subsequent event time can be generated by

$$t \leftarrow \Lambda^{-1}(\Lambda(a;z(t)) - \log(1-u))$$

where a is the last event time and u is a random number in [0, 1). Two examples show the closed form expressions for event time generation with the step and exponential link functions.

Example 4 (homogeneous Poisson baseline process)

(a) Let the link function be a step function

$$\psi(z(t)) = \begin{cases} c_1 & 0 \le t < b \\ c_2 & t \ge b. \end{cases}$$

The cumulative link function is

$$\Psi(t) = \begin{cases} c_1 t & 0 \le t < b \\ bc_1 + c_2(t-b) & t \ge b. \end{cases}$$

The cumulative intensity function with covariate $z(t)$ is

$$\Lambda(t;z(t)) = \int_0^{\Psi(t)} \lambda\,du = \begin{cases} \lambda c_1 t & 0 \le t < b \\ \lambda(c_1 b + c_2(t-b)) & t \ge b. \end{cases}$$

The time dependency associated with the covariate is effectively absorbed into the cumulative intensity function. The inverse of the cumulative intensity function is

$$\Lambda^{-1}(y) = \begin{cases} \dfrac{y}{\lambda c_1} & 0 \le y < \lambda c_1 b \\ \dfrac{y}{\lambda c_2} - \dfrac{c_1 b}{c_2} + b & y \ge \lambda c_1 b. \end{cases}$$

The general expression for the generation of the next event time, t, is closed form for $a < b$

$$t \leftarrow \begin{cases} a - \dfrac{1}{\lambda c_1} \log(1-u) & u < 1 - e^{(\lambda c_1 a - b)} \\ \dfrac{c_1}{c_2}(a-b) - \dfrac{\log(1-u)}{\lambda c_2} + b & u \ge 1 - e^{(\lambda c_1 a - b)}. \end{cases}$$

For $a \ge b$, the variate for the next event time reduces to

$$t \leftarrow a - \frac{\log(1-u)}{\lambda c_2}.$$

Figure 2 shows how the inversion of cumulative intensity technique is used to generate the subsequent event time t for a given previous event time a. For illustration, we assume the intensity $\lambda(t) = 0.1$, $c_1 = 1$, $c_2 = 2$, $a = 4$ and $b = 5$.

(b) Let the link function be an exponential function, i.e., $\psi(t) = e^{\beta t}$. The cumulative link function is

$$\Psi(t) = \frac{1}{\beta}(e^{\beta t} - 1) \qquad t \ge 0.$$

The cumulative intensity function is

$$\Lambda(t; z(t)) = \frac{\lambda}{\beta}(e^{\beta t} - 1) \qquad t \ge 0.$$

The inverse of the cumulative intensity function is

$$\Lambda^{-1}(y) = \frac{1}{\beta}\log\left(1 + \frac{y\beta}{\lambda}\right) \qquad y \ge 0.$$

The event time can be generated by the closed form expression

$$t \leftarrow \frac{1}{\beta}\log\left(e^{\beta a} - \frac{\beta}{\lambda}\log(1-u)\right).$$

Example 5 (power law baseline process)

(a) Let the link function be a step function

$$\psi(z(t)) = \begin{cases} c_1 & 0 \le t < b \\ c_2 & t \ge b. \end{cases}$$

As before, the cumulative link function is

$$\Psi(t) = \begin{cases} c_1 t & 0 \le t < b \\ bc_1 + c_2(t-b) & t \ge b. \end{cases}$$

The cumulative intensity function with covariate $z(t)$ is

$$\Lambda(t; z(t)) = \begin{cases} \nu c_1^\delta t^\delta & 0 \le t < b \\ \nu\,(c_1 b + c_2(t-b))^\delta & t \ge b. \end{cases}$$

The inverse of the cumulative intensity function is

$$\Lambda^{-1}(y) = \begin{cases} \left(\frac{y}{\nu c_1^\delta}\right)^{\frac{1}{\delta}} & y < \nu c_1^\delta b^\delta \\ \frac{1}{c_2}\left(\frac{y}{\nu}\right)^{\frac{1}{\delta}} - \frac{c_1}{c_2}\,b + b & y \ge \nu c_1^\delta b^\delta. \end{cases}$$

The general expression for variate generation is closed form for $a < b$

$$t \leftarrow \begin{cases} \left(a^\delta - \frac{1}{\nu c_1^\delta}\log(1-u)\right)^{1/\delta} & \nu c_1^\delta a^\delta - \log(1-u) < \nu (c_1 b)^\delta \\ \frac{1}{c_2}\left(\left(c_1^\delta a^\delta - \frac{1}{\nu}\log(1-u)\right)^{1/\delta} - c_1 b\right) + b & \nu c_1^\delta a^\delta - \log(1-u) \ge \nu (c_1 b)^\delta. \end{cases}$$

For $a \ge b$, the expression for the next event time reduces to

$$t \leftarrow \frac{1}{c_2}\left(\left((c_1 b + c_2(a-b))^\delta - \frac{1}{\nu}\log(1-u)\right)^{1/\delta} - c_1 b\right) + b.$$

(b) Let the link function be an exponential function, i.e., $\psi(t) = e^{\beta t}$. The cumulative link function is

$$\Psi(t) = \frac{1}{\beta}(e^{\beta t} - 1) \qquad\qquad t \geq 0.$$

The cumulative intensity function is

$$\Lambda(t; z(t)) = \nu \beta^{-\delta} (e^{\beta t} - 1)^{\delta} \qquad\qquad t \geq 0.$$

The inverse of the cumulative intensity function is

$$\Lambda^{-1}(y) = \frac{1}{\beta} \log(1 + \beta(\frac{y}{\nu})^{\frac{1}{\delta}}) \qquad\qquad y \geq 0.$$

The next event time can be generated by the closed form expression

$$t \leftarrow \frac{1}{\beta} \log(((e^{\beta a} - 1)^{\delta} - \frac{\beta^{\delta}}{\nu} \log(1-u))^{1/\delta} + 1).$$

4.2 Proportional Intensity Model

For the proportional intensity model, the cumulative intensity function with covariates is

$$\Lambda(t; z(t)) = \int_0^t \psi(z(u))\, \lambda_0(u) du$$

where ψ is the link function and λ_0 is the baseline intensity function. The subsequent event time for a single time dependent covariate, $z(t)$, can be generated by

$$t \leftarrow \Lambda^{-1}(\Lambda(a; z(t)) - \log(1-u))$$

given that the last event has occurred at time a. Three examples are used to show closed form expressions for event time generation with step and exponential link functions.

Example 6 (homogeneous Poisson baseline process)

(a) Let the link function be a step function

$$\psi(z(t)) = \begin{cases} c_1 & 0 \le t < b \\ c_2 & t \ge b. \end{cases}$$

The cumulative intensity function with covariate is

$$\Lambda(t; z(t)) = \int_0^t \psi(z(u))\lambda_0(u)du = \begin{cases} \lambda c_1 t & 0 \le t < b \\ \lambda(c_1 b + c_2(t-b)) & t \ge b. \end{cases}$$

The inverse of the cumulative intensity function is

$$\Lambda^{-1}(y) = \begin{cases} \dfrac{y}{\lambda c_1} & 0 \le y < \lambda c_1 b \\ \dfrac{y - c_1 \lambda b}{c_2 \lambda} + b & y \ge \lambda c_1 b. \end{cases}$$

Not surprisingly, the expression for variate generation is the same as that in accelerated time models (Example 4a). This is because the accelerated time and proportional intensity models have identical intensity functions under covariate effects when the baseline process is a homogeneous Poisson process.

(b) Let the link function be an exponential function, i.e., $\psi(t) = e^{\beta t}$. The cumulative intensity function is

$$\Lambda(t; z(t)) = \frac{\lambda}{\beta}(e^{\beta t} - 1) \qquad t \ge 0.$$

The event time can be generated by the same expression in Example 4b for accelerated time models.

Example 7 (log logistic baseline process)

Let the link function be a step function

$$\psi(z(t)) = \begin{cases} c_1 & 0 \le t < b \\ c_2 & t \ge b. \end{cases}$$

The cumulative intensity function with covariate $z(t)$ is $\int_0^t \psi(z(u)) \dfrac{\rho^\kappa \, \kappa u^{\kappa-1}}{1+(\rho u)^\kappa} \, du$, which yields

$$\Lambda(t; z(t)) = \begin{cases} c_1 \log(1+(\rho t)^\kappa) & 0 \le t < b \\ c_2 \log(1+(\rho t)^\kappa) + (c_1 - c_2)\log(1+(\rho b)^\kappa) & t \ge b. \end{cases}$$

The inverse of the cumulative intensity function is

$$\Lambda^{-1}(y) = \begin{cases} \dfrac{1}{\rho}(e^{y/c_1} - 1)^{1/\kappa} & y < c_1 \log(1+(\rho b)^\kappa) \\ \dfrac{1}{\rho}(e^{(y-(c_1-c_2)(\log(1+(\rho b)^\kappa)))/c_2} - 1)^{\frac{1}{\kappa}} & y \ge c_1 \log(1+(\rho b)^\kappa). \end{cases}$$

The general expression for variate generation is closed form when $a < b$

$$t \leftarrow \begin{cases} \dfrac{1}{\rho}(\dfrac{1+(\rho a)^\kappa}{(1-u)^{1/c_1}} - 1)^{1/\kappa} \\ \qquad c_1 \log(1+(\rho a)^\kappa) - \log(1-u) < c_1 \log(1+(\rho b)^\kappa) \\ \dfrac{1}{\rho}(\dfrac{(1+(\rho a)^\kappa)^{c_1/c_2}}{(1-u)^{1/c_2}(1+(\rho b)^\kappa)^{(c_1-c_2)/c_2}} - 1)^{1/\kappa} \\ \qquad c_1 \log(1+(\rho a)^\kappa) - \log(1-u) \ge c_1 \log(1+(\rho b)^\kappa). \end{cases}$$

When $a \ge b$, the expression for the next event time is

$$t \leftarrow \frac{1}{\rho}(\frac{1+(\rho a)^\kappa}{(1-u)^{1/c_2}} - 1)^{1/\kappa}.$$

When the step function has a power or exponential form, the variate generation expression is not closed form.

Example 8 (exponential power baseline process)

Let the link function be a step function

$$\psi(z(t)) = \begin{cases} c_1 & 0 \le t < b \\ c_2 & t \ge b. \end{cases}$$

The cumulative intensity function with covariate is $\int_0^t \psi(z(u)) \frac{\gamma}{\tau} (\frac{u}{\tau})^{\gamma-1} e^{(u/\tau)^\gamma} du$,

which yields

$$\Lambda(t; z(t)) = \begin{cases} c_1(e^{(\frac{t}{\tau})^\gamma} - 1) & 0 \le t < b \\ c_2\, e^{(\frac{t}{\tau})^\gamma} + (c_1 - c_2)\, e^{(\frac{b}{\tau})^\gamma} - c_1 & t \ge b. \end{cases}$$

The inverse of the cumulative intensity function is

$$\Lambda^{-1}(y) = \begin{cases} \tau(\log(\frac{y}{c_1} + 1))^{1/\gamma} & y < c_1(e^{(b/\tau)^\gamma} - 1) \\ \tau(\log(\frac{1}{c_2}(y + c_1 - (c_1 - c_2)e^{(b/\tau)^\gamma})))^{1/\gamma} & y \ge c_1(e^{(b/\tau)^\gamma} - 1). \end{cases}$$

The general expression for variate generation is closed form for $a < b$

$$t \leftarrow \begin{cases} \tau(\log(e^{(a/\tau)^\gamma} - \frac{1}{c_1}\log(1-u)))^{1/\gamma} \\ \qquad c_1(e^{(a/\tau)^\gamma} - 1) - \log(1-u) < c_1(e^{(b/\tau)^\gamma} - 1) \\ \tau(\log(\frac{1}{c_2}(c_1 e^{(a/\tau)^\gamma} - \log(1-u) - (c_1 - c_2)e^{(\frac{b}{\tau})^\gamma})))^{1/\gamma} \\ \qquad c_1(e^{(a/\tau)^\gamma} - 1) - \log(1-u) \ge c_1(e^{(b/\tau)^\gamma} - 1). \end{cases}$$

For $a \ge b$, the expression for the next event time is

$$t \leftarrow \tau(\log(e^{(\frac{a}{\tau})^\gamma} - \frac{1}{c_2}\log(1-u)))^{1/\gamma}.$$

When the step function has a power or exponential form, the expression is not closed form.

Two tables are presented to summarize the results in Section 4. Table 4 indicates whether closed form expressions are available for lifetime (first event time) generation when different baseline distributions and covariate functions are assumed. Table 5 shows whether closed form expressions are available for event times generation when different baseline processes and covariate functions are assumed.

	AL			PH		
Baseline: link:	step	exponential	power	step	exponential	power
Exponential	C	C	C	C	C	C
Weibull	C	C	C	C	N	C
Log logistic	C	C	C	C	N	N
Exponential power	C	C	C	C	N	N

Table 4. Summary of lifetime variate generation with time dependent covariates.

	AT			PI		
Baseline: link:	step	exponential	power	step	exponential	power
HPP	C (4a)	C (4b)	C	C (6a)	C (6b)	C
Power	C (5a)	C (5b)	C	C	N	N
Log logistic	C	C	C	C (7)	N (9)	N
Exponential power	C	C	C	C (8)	N	N

Table 5. Summary of event times variate generation with time dependent covariates.

In each box, we use "C" to denote that a closed form expression for the event time generation is available. The number in the parenthesis indicates the example that includes the derivation of the expression. We use "N" (not closed form) when a closed form expression is not available.

When there is no closed form expression using inversion of the cumulative intensity, more complicated variate generation techniques, such as thinning, can be considered for generating event times for the models. Example 9 illustrates the use of thinning, assuming the log logistic baseline process and the exponential link function.

Example 9 (baseline process with log logistic intensity)

There is no closed form expression for variate generation by inversion when the baseline process is log logistic (with shape parameter $\kappa \geq 1$) and the link function is exponential. The baseline cumulative intensity and intensity functions are

$$\Lambda_0(t) = \log\left(1 + (\rho t)^\kappa\right) \qquad\qquad \lambda_0(t) = \frac{\rho^\kappa \kappa t^{\kappa - 1}}{1 + (\rho t)^\kappa}.$$

The link function $\psi(z(t))$ is

$$\psi(z(t)) = e^{\beta t}.$$

where β is the regression coefficient. The intensity function under the proportional intensity assumption is

$$\lambda(t; z(t)) = \frac{\rho^\kappa \kappa t^{\kappa - 1} e^{\beta t}}{1 + (\rho t)^\kappa}.$$

In this example, thinning can be used since the baseline intensity function has an upside down bathtub shape when $\kappa \geq 1$ and the intensity function under covariate effects $\lambda(t; z(t))$ remains finite when the regression coefficient is negative. When thinning is used, a majorizing function needs to be defined. One simple, albeit computationally inefficient, choice for the majorizing function is the maximum value of the intensity function $\lambda(t; z(t))$. This typically requires much more CPU time since the probability of a rejection is fairly high for most time values.

To generate event times for the baseline process, the majorizing function λ_0^* is required. The point that maximizes the baseline intensity can be obtained by equating the first derivative of $\lambda_0(t)$ with respect to t to zero. The maximum baseline intensity is at $t_0^* = \dfrac{(\kappa - 1)^{1/\kappa}}{\rho}$, where superscript ’*’ denotes the maximum point. The maximum of baseline intensity $\lambda_0(t_0^*) = \lambda_0^* = \rho(\kappa - 1)^{1-1/\kappa}$ is therefore used as a majorizing (constant) function to generate event time for the baseline process. For example, when $\kappa = 2$ and $\rho = 0.5$, the majorizing function is $\lambda_0^* = 0.5$ and $t_0^* = 2.0$.

For the process under covariate effects, the maximum value of $\lambda(t; z(t))$ is obtained by equating the first derivative of $\lambda(t; z(t))$ with respect to t to zero.

$$\beta \rho^{\kappa} t^{\kappa+1} - \rho^{\kappa} t^{\kappa} + \beta t + \kappa - 1 = 0.$$

In general, there is not a closed form solution for t_c^* which solves the equation. One specific value for the shape parameter, $\kappa = 2$, is used here to illustrate the algorithm. In this case, the equation reduces to the cubic equation

$$\beta \rho^2 t^3 - \rho^2 t^2 + \beta t + 1 = 0.$$

To illustrate, we further assume that $\rho = 0.5$ and $\beta = -0.1$. The maximum of $\lambda(t; z(t))$ is obtained by solving a nonlinear equation with the bisection method. The solution is $t_c^* = 1.6868$, where subscript 'c' denotes the condition under covariate effects. The corresponding maximum value $\lambda(t_c^*; z(t_c^*)) = \lambda_c^* = 0.4164$ can be used as the majorizing function to generate event times under covariate effects. Figure 3 shows the two intensity functions and their majorizing functions, λ_0^* and λ_c^*, where $\kappa = 2$, $\rho = 0.5$ and $\beta = -0.1$ are assumed.

To illustrate the variate generation algorithm for the process under covariate effects using thinning, the next event time t_{next} is generated given that the last event has occurred at time a.

1. [Set the current event time.] $t \leftarrow a$

2. [Generate random numbers.] $u_1, u_2 \sim IID\ U(0,1)$

3. [Determine the next event time.] $t \leftarrow t + (\frac{-1}{\lambda_c^*} \log(1 - u_1))$

4. [Thinning.] if $u_2 > \lambda(t; z(t)) / \lambda_c^*$ go to 2

5. [Set the next event time.] $t_{next} \leftarrow t$

Figure 4 depicts the event time generation by thinning, where a denotes the given last event time, t_{next} is the next event time and t's are the rejected event times by thinning in step 4. To generate the event time for baseline process, the majorizing function is replaced by λ_0^* and $\lambda(t; z(t))$ is replaced by $\lambda_0(t)$ in steps 3 and 4.

References

Allison, P.D. (1984). *Event History Analysis: Regression for Longitudinal Event Data.* Beverly Hills, California, SAGE Publications Ltd.

Andersen, P.K. and Gill, R.D. (1982). Cox's Regression Model for Counting Processes: A Large Sample Study. *The Annals of Statistics* (10), 1100-1120.

Cinlar, E. (1975). *Introduction to Stochastic Processes.* Englewood Cliffs, New Jersey, Prentice-Hall.

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

Dale, C.J. (1985). Application of the Proportional Hazards Model in the Reliability Field. *Reliability Engineering* (10), 1-14.

Devroye, L. (1986). *Non-Uniform Random Variate Generation.* New York, Springer-Verlag.

Fishman, G.S. (1978). *Principles of Discrete Event Simulation.* New York, John Wiley.

Fishman, G.S. and Kao, E.P.C. (1977). A Procedure for Generating Time-Dependent Arrivals for Queueing Simulations. *Naval Research Logistics Quarterly* (24) **4,** 661-666.

Hoffman, R.G. (1985). Simulated Temporal Models for Proportional Hazards Models with Time Varying Covariates. American Statistical Association, Statistical Computing Section, 390-395.

Kalbfleisch, J.D. and McIntosh, A.A. (1977). Efficiency in Survival Distributions with Time-Dependent Covariables. *Biometrika* (64) **1,** 47-50.

Kaminsky, F.C. and Rumpf, D.L. (1977). Simulating Nonstationary Poisson Processes: A Comparison of Alternatives Including the Correct Approach. *Simulation* (29) **1,** 17-20.

Karr, A.E. (1986). *Point Processes and Their Statistical Inference.* New York, Marcel Dekker.

Klein, R.W. and Roberts, S.D. (1984). A Time-Varying Poisson Arrival Process

Generator. *Simulation* (43) **4,** 193-195.

Lavenberg, S.S. (1983). *Computer Performance Modeling Handbook.* Orlando, Florida, Academic Press.

Law, A. and Kelton, D. (1991). *Simulation Modeling and Analysis, Second Edition.* New York, McGraw-Hill.

Lawless, J.F. (1987). Regression Methods for Poisson Process Data. *Journal of the American Statistical Association* (82) **399,** 808-815.

Lee, S., Wilson, J.R. and Crawford, M.M. (1991). Modeling and Simulation of A Nonhomogeneous Poisson Process with Cyclic Features. *Communications in Statistics-Simulation and Computation* (20) **2&3,** 777-809.

Leemis, L.M. (1987). Variate Generation for Accelerated Life and Proportional Hazards Models. *Operations Research* (35) **6,** 892-894.

Leemis, L.M. (1991). Nonparametric Estimation of the Cumulative Intensity Function for A Nonhomogeneous Poisson Process. *Management Science* (37) **7,** 886-900.

Leemis, L.M., Shih, L. and Reynertson, K. (1990). Variate Generation for Accelerated Life and Proportional Hazards Models with Time Dependent Covariates. *Statistics & Probability Letters* (10) **4,** 335-339.

Lewis, P.A.W. and Shedler, G.S. (1976). Simulation of Nonhomogeneous Poisson Processes with Log Linear Rate Function. *Biometrika* (63) **3,** 501-505.

Lewis, P.A.W. and Shedler, G.S. (1979a). Simulation of Nonhomogeneous Poisson Processes with Degree-Two Exponential Polynomial Rate Function. *Operations Research* (27) **5,** 1026-1040.

Lewis, P.A.W. and Shedler, G.S. (1979b). Simulation of Nonhomogeneous Poisson Processes By Thinning. *Naval Research Logistics Quarterly* (26), 403-413.

Ogata, Y. (1981). On Lewis' Simulation Method for Point Processes. *IEEE Transactions on Information Theory* (27) **1,** 23-31.

Petersen, T. (1986a). Fitting Parametric Survival Models with Time-Dependent Covariates. *Applied Statistics* (35) **3,** 281-288.

Petersen, T. (1986b), Estimating Fully Parametric Hazard Rate Models with Time-

Dependent Covariates, *Sociological Methods & Research* (14) **3,** 219-246.

Prentice, R.L. and Kalbfleisch, J.D. (1979). Hazard Rate Models with Covariates. *Biometrics* (35), 25-39.

Prentice, R.L., Williams, B.J. and Peterson, A.V. (1981). On the Regression Analysis of Multivariate Failure Time Data. *Biometrika* (68) **2,** 373-379.

Ross, S.M. (1990). *A Course in Simulation.* New York, Macmillan.

Smith, J.R. (1987). Reliability Analysis by Simulation, ASQC 1987 Annual Technical Conference Transactions, 654-662.