



Dipartimento di Statistica
"Giuseppe Parenti"

Dipartimento di Statistica "G. Parenti" – Viale Morgagni 59 – 50134 Firenze – www.ds.unifi.it

W O R K I N G P A P E R 2 0 0 2 / 0 7

Graphical Duration Models

Anna Gottard



Università degli Studi
di Firenze

Statistics

GRAPHICAL DURATION MODELS

Anna Gottard

Department of Statistics, University of Florence, Italy

ABSTRACT. Graphical models use graphs to represent conditional independence relationships among random variables. In this paper, the possibility of defining a special type of graphs including nodes to represent point processes has been explored: this could be relevant to the analysis of *event history* data, i.e. of data consisting of times and types of observed events, together with some explanatory variables observed on a number of individuals; survival analysis and duration models are particular cases. The idea consists of representing a whole process by a single node. The resulting models will be called graphical duration models.

Key words: chain graphs, duration models, event history analysis, graphical model, marked point processes.

1. Introduction

Event history, duration data or generic time-evolving processes arise frequently in many fields, such as demography, sociology, epidemiology and so on. A routine objective is simply the study of dependence of time to events on some explanatory variables, but situations of complex data structure are frequently encountered: different event histories can interact with one other and depend on a large number of potentially explanatory variables, both time-dependent and time constant. Much work has been developed in literature focusing on interdependence between times to events, according to the specific model formulation and types of data. See for example Cox (1972) for bivariate survival times as point processes, van den Berg (1997) for bivariate hazard rate models, and Basu and Sun (1997) for multivariate failure rates based on Cox models, among others. For the general framework of event history analysis, there are two main approaches: the so-called *causal approach* of Blossfeld *et al.* (1995), and the *system approach*, utilized by Tuma and Hannan (1984), but firstly described by Cox and Lewis (1972) and Aalen *et al.* (1980). The system approach consists of representing a joint process $\{X(t, m_X), Y(t, m_Y)\}$ as a marked process $W(t, m_W)$, taking values on the product state space $\mathcal{S}_W = \mathcal{S}_X \times \mathcal{S}_Y$. On one hand, this approach has the advantage of being applicable to almost all type of point processes and of not requiring strong assumptions as the *causal* approach. On the other, it is complex to implement and, in particular, to interpret. In the framework of random variables, the existence of both symmetric and asymmetric relationships is cleverly investigated by graphical models, using the so called *chain graphs* (Lauritzen

and Wermuth, 1989, Whittaker, 1990, Cox and Wermuth, 1996 and Lauritzen, 1996), a particular class of graphical Markov models.

This paper explores the conditions under which chain graphs could be able to identify the set of conditional independence relationships for a class of models for event history or duration data, considered as marked point processes (MPPs), (Arjas, 1989). The basic proposal put forward here, in accordance with chain graph theory, consists of representing a whole marked point process by a single node and conditional independence by a missing edge. The class of models for event history or duration data representable by chain graphs are here called *graphical duration models*.

This solution can be advantageous for both graphical models, whose application field is so extended, and duration models, which gain a valid instrument of analysis and interpretation. In particular, chain graphs could be helpful for the deduction of conditional independence structure between variables and event histories; moreover, they suggest useful ways to simplify the inferential procedure and facilitate the intuitive representation of such complex models.

Former contribution on the topic is Didelez (2001), where directed graphs for Markov processes have been developed: she defines a kind of directed graph, called local independence graphs, where a node represent a Markov process and an arrow local dependence. The present paper generalizes these results by taking into account graphs which can include both marked point processes and random variables, and both directed and undirected edges. In addition, the general class of marked point process allowing for Doob-Meyer decomposition is considered.

The rest of the paper is organized as follows: in the next section, the main definitions of conditional independence between MPPs and between variables and MPPs are discussed; section 3 presents the formal definition of graphical duration models; as an illustration, a simple example of application of graphical duration models to real data is shown in section 4. Last section is dedicated to some brief concluding remarks.

2. On conditional independence for MPPs

In this section, the definitions for conditional independence involving marked point processes and random variables are reviewed. First of all, let us consider the definition of a MPP $X(t, m_X)$ independence of a random variable, Z . The reverse definition of Z on $X(t, m_X)$ is not analyzed, as it is here assumed that a random variable, which is constant with respect to time, can not depend on something which is indexed by time.

One can say that a MPP $X(t, m_X)$ is independent ($\perp\!\!\!\perp$) of a random variable Z if its probability distribution is not a function of Z . In other words, if $f_X(t) = f_X(t_1, m_1, t_2, m_2, \dots)$ is the density function of the whole MPP $X(t, m_X)$,

$$X(t, m_X) \perp\!\!\!\perp Z \text{ if and only if } f_X(t | Z) = f_X(t) \quad (1)$$

for all $t > 0$, or, equivalently, considering mark specific hazard functions,

$$h_{m_X}(t | Z) = h_{m_X}(t) \text{ for all } t > 0 \text{ and } m_X \in \mathcal{M}_X.$$

This equivalence derives directly from the relationship between the density, hazard and survival functions of a MPP. Previous definition can be generalized to the case of a MPP and two random variables: for all $t > 0$,

$$X(t, m_X) \perp\!\!\!\perp Z_2 | Z_1 \text{ if and only if } f_X(t | Z_1, Z_2) = f_X(t | Z_1). \quad (2)$$

Once again, condition (2) is equivalent to $h_m(t | Z_1, Z_2) = h_{m_X}(t | Z_1)$ for all $t > 0$ and $m_X \in \mathcal{M}_X$.

To define conditional independence between MPPs, to fix the meaning of dependence and independence between them is needed: this may involve further differences with respect to random variables definitions. One of the first attempts in this issue is due to Schweder (1970) who gave the definition of *local independence* between Markov processes; this concept has been then generalized by Aalen (1987) for all processes allowing for Doob-Meyer decomposition. The main result on the topic is the existence of two kinds of independence between stochastic processes: *local independence*, related to a one-sided influence, and stochastic independence, that is a mutual local independence. In particular, the definition of local independence is the following:

A process $Y(t)$ is said to be locally independent of a process $X(t)$ over some time interval, if the risks of changes in $Y(t)$ are independent of the value of $X(t)$ for all times in the interval. In addition, $Y(t)$ is said to be locally dependent on $X(t)$ if it is not locally independent.

Thus, $Y(t, m_Y)$ can be locally independent of $X(t, m_X)$, while $X(t, m_X)$ is locally dependent on $Y(t, m_Y)$; it implies that the condition of local independence is not symmetric. Still, generalizing Schweder's theorem 3, Aalen assesses that two processes $X(t, m_X)$ and $Y(t, m_Y)$ are mutually locally independent over a time interval if and only if they are also stochastically independent, which is a symmetric definition. Here and in the following, the notation $\perp\!\!\!\perp$, will be used for local independence and $\perp\!\!\!\perp$ for stochastic independence.

As pointed out by Courgeau and Lelièvre (1992), by system approach it is possible to verify the existence of stochastic or local independence between event histories. The following proposition parallels their results in terms of MPPs.

PROPOSITION 1. *Let $X(t, m_X)$, $t > 0$, $m_X \in \mathcal{M}_X$ and $Y(t, m_Y)$, $t \geq 0$, $m_Y \in \mathcal{M}_Y$ be two marked point processes, taking values on \mathcal{S}_X and \mathcal{S}_Y , respectively; let $W(t, m_W)$, $t > 0$, $m_W \in \mathcal{M}_W$ be the marked process with values on the product-space of states $\mathcal{S}_W = \mathcal{S}_X \times \mathcal{S}_Y$. If there exists a proper partition of marks space $\mathcal{M}_W = (\mathcal{M}_X^W, \mathcal{M}_Y^W)$ such that,*

$$\mathcal{M}_X^W \equiv \mathcal{M}_X \text{ and } \mathcal{M}_Y^W \equiv \mathcal{M}_Y \quad (3)$$

then $X(t, m_X)$ and $Y(t, m_Y)$ are stochastically independent.

Otherwise, if the partition of marks space $\mathcal{M}_W = (\mathcal{M}_X^W, \mathcal{M}_Y^W)$ exists and is such that

$$\mathcal{M}_X^W \neq \mathcal{M}_X, \mathcal{M}_Y^W \equiv \mathcal{M}_Y \quad (4)$$

then $Y(t, m_Y)$ is locally independent of $X(t, m_X)$. Similarly for the reverse.

The proof is sketched in the appendix. To illustrate Proposition 1, a simple case can be considered: let $X(t, m_X)$ and $Y(t, m_Y)$ be two MPPs with state spaces $\mathcal{S}_X = \{0, 1\}$ and $\mathcal{S}_Y = \{0, 1\}$ respectively, that do not change state in the same instant of time. By the system approach, the joint process $\{X(t, m_X), Y(t, m_Y)\}$ is represented as a marked point process $W(t, m_W)$, taking values on the product state space $\mathcal{S}_W = \mathcal{S}_X \times \mathcal{S}_Y = \{(0, 0), (0, 1), (1, 0), (1, 1)\}$. Its marks space is then composed by the following elements:

$$\begin{array}{llll} m_W = 1 & \text{for} & (0, 0) & \longrightarrow & (0, 1) & m_W = 2 & \text{for} & (0, 0) & \longrightarrow & (1, 0) \\ m_W = 3 & \text{for} & (0, 1) & \longrightarrow & (0, 0) & m_W = 4 & \text{for} & (0, 1) & \longrightarrow & (1, 1) \\ m_W = 5 & \text{for} & (1, 0) & \longrightarrow & (0, 0) & m_W = 6 & \text{for} & (1, 0) & \longrightarrow & (1, 1) \\ m_W = 7 & \text{for} & (1, 1) & \longrightarrow & (1, 0) & m_W = 8 & \text{for} & (1, 1) & \longrightarrow & (0, 1). \end{array}$$

In this case, the existence of the partition of $\mathcal{M}_W = \{1, 2, 3, 4, 5, 6, 7, 8\}$ in $\mathcal{M}_X^W = \{2, 4, 5, 8\}$ and $\mathcal{M}_Y^W = \{1, 3, 6, 7\}$ can be easily noted. Therefore, according to condition (3), $X(t, m_X) \perp\!\!\!\perp Y(t, m_Y)$ if

$$h_1^W(t) = h_6^W(t) \quad h_2^W(t) = h_4^W(t) \quad h_3^W(t) = h_7^W(t) \quad h_5^W(t) = h_8^W(t).$$

where $h_m^W(t)$ is the mark m specific hazard function of $W(t, m_W)$. Still, by condition (4) $Y(t, m_Y) \perp\!\!\!\perp X(t, m_X)$ if $h_1^W(t) = h_6^W(t)$ and $h_5^W(t) = h_8^W(t)$, while $h_2^W(t) \neq h_4^W(t)$ and $h_3^W(t) \neq h_7^W(t)$. Similarly for the other situations.

Remark 1. Let Z be a vector of explanatory random variables of dimension $q \geq 1$. If proposition 1 is considered with respect to conditional distribution of the processes, given Z , then conditional independence and conditional local dependence can be assessed.

Remark 2. If two MPPs $X(t, m_X)$ and $Y(t, m_Y)$ have simultaneous events, then the partition $\mathcal{M}_W = (\mathcal{M}_X^W, \mathcal{M}_Y^W)$ does not exist and $X(t, m_X)$ and $Y(t, m_Y)$ cannot be stochastic or local independent. Regularity is therefore a necessary condition for stochastic and local independence and conditional independence between two MPPs. Besides, it is a necessary and sufficient condition for the existence of the partition of mark space requested in Proposition 1.

Remark 3. Assuming a parametric statistical model for the MPPs, the proposition gives a way to assess marginal or conditional independence between two (or more) marked point processes, indicating that both the independence and the local dependence cases can be modelled as a reduced version of the mutual dependence model.

3. Definition of graphical duration models

We are now in a position to define a graphical model for event history data. In addition to random variables, which are usually represented as circles or dots, a MPPs can be represented by a node in a graph. The resulting graph is $\mathcal{G} = (V, E)$, where V , the set of nodes, can be partitioned in (Γ, Δ, Π) , with nodes in Γ representing continuous variables (\circ), those in Δ discrete variables (\bullet) and those in Π MPPs ($\overline{\circ}$). Further, E , the set of edges, is a set of ordered pairs of distinct nodes, representing the presence of directed (arrow) or undirected (line) edges between two nodes: if $u, v \in (\Gamma \cup \Delta) \vee \Pi$, there is a line between u and v , denoted $u - v$, if both (u, v) and $(v, u) \in E$, while there is an arrow pointing to v , denoted $u \rightarrow v$, if $(u, v) \in E$, but $(v, u) \notin E$. If $u \in (\Gamma \cup \Delta)$ and $v \in \Pi$, then either there is an arrow pointing to v or both (u, v) and $(v, u) \notin E$. By construction, this kind of graphs can be always decomposed into two subgraphs $\mathcal{G}_{\Gamma \cup \Delta}$ and \mathcal{G}_{Π} , where a node in $\mathcal{G}_{\Gamma \cup \Delta}$ cannot be a child or a neighbor of a node in \mathcal{G}_{Π} . Moreover, nodes in V can be also partitioned into an ordered sequence of blocks, identifying the chain components, so that variables placed in a same block have to be treated on an equal footing. Each block can contain only processes or variables; blocks containing variables must always precede blocks with processes.

A graphical duration model for V is specified by assuming that the joint distribution of variables and MPPs in V belongs to the set of distributions which obey the chain graph Markov property with respect to \mathcal{G} . The absence of connection between two variables has the classical meaning of conditional independence; the absence of an arrow between a variable and a process implies that the process is conditionally independent of the variable, given all the remaining processes in the box and all variables and processes in the preceding boxes. An arrow between two processes means conditional local independence, while the absence of connection stochastic conditional independence. The chain graph Markov property manifests itself through the existence of a recursive factorization of the joint distribution on V . In this case, specifically, the outer factorization is more complex and involves some passages. Firstly, the partition of \mathcal{G} in $\mathcal{G}_{\Gamma \cup \Delta}$ and \mathcal{G}_{Π} and the order of blocks implies that the joint density f on \mathcal{G} can be always factorized in the form

$$f(x) = f(x_{\Gamma \cup \Delta}) \cdot f(x_{\Pi} | Z_{\Gamma \cup \Delta}) \quad (5)$$

where x_A denotes a configuration $(x_v)_{v \in A}$ of subset of random quantities $A \subseteq V$. Secondly, the joint density function of variables in $\Gamma \cup \Delta$ factorizes (Lauritzen, 1996) in the form

$$f(x_{\Gamma \cup \Delta}) = \prod_{\tau \in \mathcal{C}_{\Gamma \cup \Delta}} f(x_{\tau} | x_{pa(\tau)}) \quad (6)$$

with $\mathcal{C}_{\Gamma \cup \Delta}$ denoting the set of chain component of $\mathcal{G}_{\Gamma \cup \Delta}$ and $x_{pa(\tau)}$ the set of parents of nodes in τ . Then, factors in (6) can be further factorized as

$$f(x_{\tau} | x_{pa(\tau)}) = \prod_{\tau \in a} \phi_a(x)$$

where a are all the cliques in $\tau \cup pa(\tau) \subseteq \mathcal{C}_{\Gamma \cup \Delta}$. Finally, if Π consists of q MPPs, the second factor in (5) is specified by the distribution of the MPP $W(t, m_W)$ taking values on the product state space. If there exists a partition in $\mathcal{M}_W = \cup_{i=1}^q \mathcal{M}_{X_i | \times X \setminus X_i}$, where $\times X \setminus X_i$ is the product state space of all point processes except $X_i(t)$, then the density function factorization is

$$f(x_\Pi | Z_{\Gamma \cup \Delta}) = \prod_{i=1}^q \prod_{\mathcal{M}_{X_i | \times ne(X_i)}} f(X_i(t, m_i) | pa(X_i)) \quad (7)$$

where $\times ne(X_i)$ denotes the product space of states of point processes whose nodes are *neighbors* of X_i . Therefore the last part of factorization occurs with respect to marks space. Because of the restriction to directed edges from variables to MPPs, the entire graph \mathcal{G} cannot be moralized.

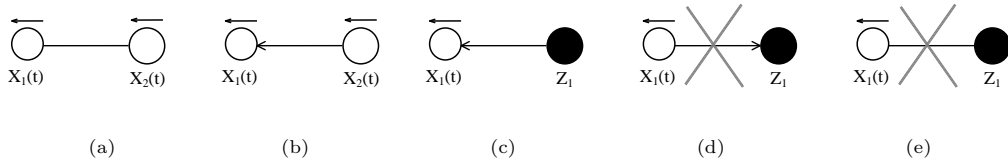


Fig. 1. Example of graphs including point processes.

Figure 1(a) illustrates the simplest graph consisting of two nodes representing two point processes: the connection by an undirected edge implies that mutual dependence is assessed; in this case the two processes are mutually locally dependent. Figure 1(b) depicts a situation of unilateral dependence: the graph, where $X_1(t, m_1)$ and $X_2(t, m_2)$ are connected by an arrow, suggests that $X_2(t, m_2)$ is locally independent of $X_1(t, m_1)$, while $X_1(t, m_1)$ is locally dependent on $X_2(t, m_2)$. Any edge between a process and a random variable must be a directed edge with the arrow pointing from the variable to the process (fig. 1(c) and (d)). Undirected lines or arrows from processes to variables are forbidden (fig. 1(e)) so that undirected graph consists only of random variables or point processes, namely MPPs and random variables cannot be represented in a same undirected graph or in a same chain component.

A situation of mutual dependence between three MPPs is described a complete undirected graph (fig. 2(a)), while the graph illustrated in figure 2 (b), where $X_1(t, m_{X_1})$ and $X_2(t, m_{X_2})$ are not connected, implies that $X_1(t, m_{X_1}) \perp\!\!\!\perp X_2(t, m_{X_2}) | X_3(t, m_{X_3})$. The arrows pointing from $X_2(t, m_{X_2})$ to $X_1(t, m_{X_1})$ in figure 2 (c), finally, indicates $X_2(t, m_{X_2}) \perp\!\!\!\perp X_1(t, m_{X_1}) | X_3(t, m_{X_3})$, while the missing connection between $X_1(t, m_{X_1})$ and $X_3(t, m_{X_3})$ implies that $X_1(t, m_{X_1}) \perp\!\!\!\perp X_3(t, m_{X_3}) | X_2(t, m_{X_2})$.

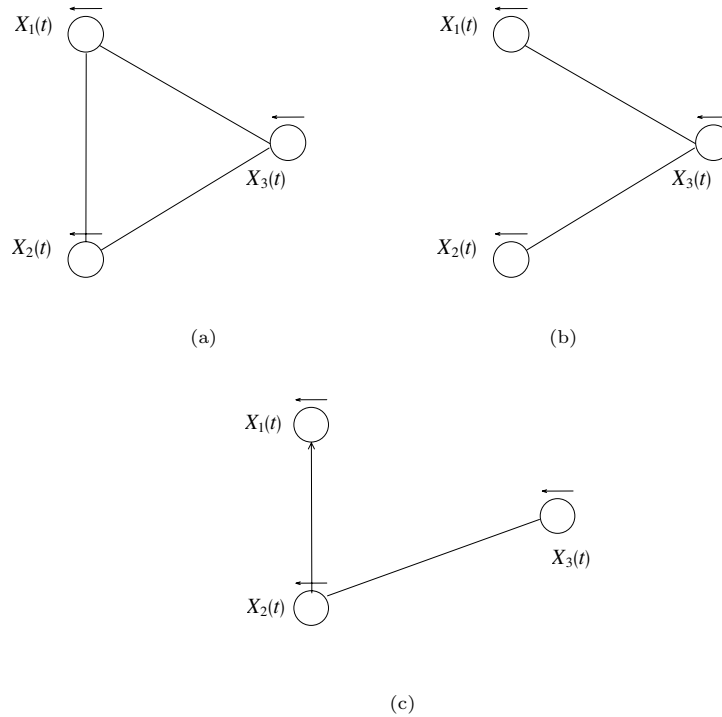


Fig. 2. Example of graphs including three point processes.

4. An example with real data

In this section, a graphical duration model is applied to study the relationship between fertility and female employment. The data used are from INF-2 survey about control and expectations of fertility in Italy (De Sandre *et al.*, 1997): this is a retrospective study carried out in the period November 1995 – January 1996 on a sample of 4824 women of age 20-49 years. The data set collect information on partnerships, fertility, employment and socio-demographic characteristics. In this example, it seems interesting to investigate the relationship between fertility (motherhood) and female labor force participation viewed as two interdependent marked process, possibly influenced by some original family characteristics, for the cohort of Italian women born in 1946-50. The model with a larger set of explanatory variables is more complex to illustrate, but the basic concepts remain the same (see for example Gottard, 1999).

Hence, let $X(t, m_X)$ denote the motherhood process, taking values on $\mathcal{S}_X = \{0, 1\}$, where 0 is the state “to have no child” and 1 the state “to be mother”; its marks space \mathcal{M}_X consists of only one element, because “to become mother” is the only allowed transition, thus $X(t, m_X)$ is a single-point process. Likewise, let $Y(t, m_Y)$ be the process describing labor-force participation, taking values on $\mathcal{S}_Y = \{0, 1\}$, where 0 is the state “to be unemployed” and 1 is the state “to be employed”. Its marks space is $\mathcal{M}_Y = \{01, 10\}$,

where $m = 01$ indicates the event “to find a job” and $m = 10$ “to leave a job”. Among the many variables describing the original family characteristics, three variables has been included: Z_1 indicating the father’s education level (1 = Secondary or higher, 0 = otherwise), Z_2 for mother’s working experience (1 = always or nearly always employed, 0 = otherwise) and Z_3 for the dimension of the original family (1 = large family, that is more then two brothers and/or sisters, 0 = otherwise).

The set V is so composed of five nodes: two MPPs and three discrete variables. Nodes are ordered into two blocks: the first block (on the left) contains the explanatory variables, whereas the second block contains the two MPPs. According to (5) the joint distribution of the random quantities can be so factorized $f(x_t, y_t, z_1, z_2, z_3) = f(x_t, y_t | z_1, z_2, z_3)f(z_1, z_2, z_3)$ and hence the likelihood function

$$\mathcal{L}(\theta; x_t, y_t, z_1, z_2, z_3) = \mathcal{L}_1(\theta_1; x_t, y_t, z_1, z_2, z_3) \cdot \mathcal{L}_2(\theta_2; z_1, z_2, z_3), \quad (8)$$

where θ is the vector of unknown real parameters, with $\theta = \theta_1 \cup \theta_2$ and $\theta_1 \cap \theta_2 = \emptyset$. Its maximum can be achieved separately for the two factors, therefore the two components of (8) have been handled separately. About the first factor, let $W(t, m_W)$ be the marked process taking values on the product state space $\mathcal{S}_W = \mathcal{S}_X \times \mathcal{S}_Y = \{(0, 0), (0, 1), (1, 0), (1, 1)\}$; its marks space \mathcal{M}_W is composed of six elements:

$$\begin{array}{llll} m_W = 1 & \text{for} & (0, 0) & \longrightarrow & (0, 1) & & m_W = 2 & \text{for} & (0, 0) & \longrightarrow & (1, 0) \\ m_W = 3 & \text{for} & (0, 1) & \longrightarrow & (0, 0) & & m_W = 4 & \text{for} & (0, 1) & \longrightarrow & (1, 1) \\ m_W = 5 & \text{for} & (1, 0) & \longrightarrow & (1, 1) & & m_W = 6 & \text{for} & (1, 1) & \longrightarrow & (1, 0). \end{array}$$

Because of the two MPPs do not jump in the same instant of time, \mathcal{M}_W can be partitioned in $\mathcal{M}_X^W = \{2, 4\}$ and $\mathcal{M}_Y^W = \{1, 3, 5, 6\}$, and proposition 1 can be applied. Moreover, it is legitimate to assume that processes under study and censoring process are independent, that is the censure is uninformative. Hypothesizing that both the processes have the same probabilistic model, in particular a log-logistic distribution, the mark specific hazard function has the form

$$h_m(t|Z) = \frac{\lambda_m^{\alpha_m} \alpha_m t^{\alpha_m - 1}}{1 + (\lambda_m t)^{\alpha_m}} \quad (9)$$

where $\lambda_m = \exp(Z' \beta_m)$ and Z is the vector of explanatory variables (Z_1, Z_2, Z_3). The log-logistic distribution characteristics assure a great flexibility, non-proportional risks and a not monotone hazard function. Starting from a complete graph, the first factor of (8) is

$$\mathcal{L}_1 = \prod_{i=1}^n \prod_{j=1}^{n_i} \prod_{m \in \mathcal{M}_W} \left\{ h_m(t_{ij} | \mathfrak{S}_{t_{ij}^-}^W, Z_i)^{\delta_{ij}} \cdot \bar{S}_m(t_{ij} | \mathfrak{S}_{t_{ij}^-}^W, Z_i) \right\}^{\zeta_{ij}^m}, \quad (10)$$

where δ_{ij} and ζ_{ij}^m are censure and mark indicator; for notational convenience, $\bar{S}_m(t_{ij})$ indicates the product of specific survival functions relative to risks in competition with m . The history of process, $\mathfrak{S}_{t_{ij}^-}^W$, has been summarized by the sum of past durations.

Table 1. Maximum likelihood estimates of full model

<i>Explanatory Variables</i>	<i>Marks</i>					
	1	2	3	4	5	6
Constant	-3.1200 (0.0361)	-3.4330 (0.0388)	-1.7860 (0.1794)	-3.2888 (0.1719)	-4.7668 (0.3080)	-2.2403 (0.3163)
Father's education level	-0.1020 (0.0697)	-0.1945 (0.0820)	0.1040 (0.1517)	-0.2047 (0.1477)	-0.4181 (0.3580)	-0.7805 (0.2813)
Mother's working experience	0.1169 (0.0478)	-0.0552 (0.0520)	0.3083 (0.0955)	0.1856 (0.0924)	0.3453 (0.1795)	0.1129 (0.1433)
Dimension of the original family	-0.0306 (0.0453)	0.1107 (0.0446)	0.0787 (0.0940)	-0.0609 (0.0912)	-0.6120 (0.1724)	-0.0844 (0.1410)
Internal History	0.0909 (0.0019)	0.0859 (0.0019)	-0.0143 (0.0081)	0.0448 (0.0073)	0.0680 (0.0085)	-0.0133 (0.0093)
α^2	1.5954 (0.0270)	1.8124 (0.0408)	1.2876 (0.0269)	1.4351 (0.0358)	0.9700 (0.0242)	1.1130 (0.0285)

As far as the second factor of (8) is concerned, the nature of variables in analysis suggests the formulation of graphical log-linear model, corresponding to an undirected sub-graph. Calling $p(z)$, with $z = (z_1, z_2, z_3)$, the cell probabilities of the three-way table, a complete undirected sub-graph leads to the log-linear expansion

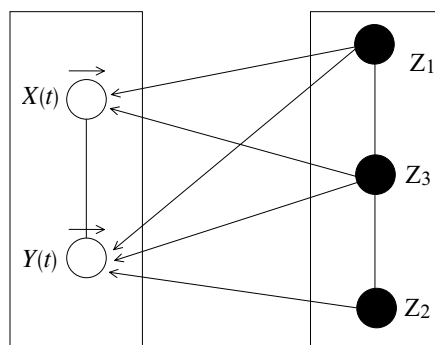
$$\begin{aligned} \log(p(z)) = & \mu_0 + \mu_1(z_1) + \mu_2(z_2) + \mu_3(z_3) + \\ & + \mu_{12}(z_1, z_2) + \mu_{13}(z_1, z_3) + \mu_{23}(z_2, z_3) + \mu_{123}(z_1, z_2, z_3) \end{aligned}$$

Results are achieved using an *ad hoc* program written in GAUSS, for the first sub-graph and MIM (Edwards, 2000) for the second one. The maximum likelihood estimates for the complete sub-graph of first factor are illustrated in table 1.

Likelihood ratio test (LR) can be used to simplify the graph. Denoting $\vartheta_i = (\beta_{0i}, \dots, \beta_{4i}, \alpha_i)$ the entire set of parameters relative to mark i specific hazard function of $W(t, m_W)$, the model for conditional independence between $X(t, m_X)$ and $Y(t, m_Y)$ is formulated putting $\vartheta_1 = \vartheta_5$, $\vartheta_2 = \vartheta_4$ and $\vartheta_3 = \vartheta_6$. Furthermore, the model for conditional local independence, with $X(t, m_X) \perp\!\!\!\perp Y(t, m_Y) \mid Z$, has $\vartheta_2 = \vartheta_4$, while the model with $Y(t, m_Y) \perp\!\!\!\perp X(t, m_X) \mid Z$ has constrained on $\vartheta_1 = \vartheta_5$ and $\vartheta_3 = \vartheta_6$. The reduced model for the first factor assumes mutual dependence between the MPPs and $X(t, m_X) \perp\!\!\!\perp Z_2 \mid Y(t, m_Y), Z_1, Z_3$; that for the second factor put $Z_1 \perp\!\!\!\perp Z_2 \mid Z_3$. Results are given in table 2: with a level of significance fixed at 0.01, the selected model is the

Table 2. Comparison of fitted models

<i>Model</i>	<i>log-lik</i>	<i>n par.</i>	<i>LR</i>	<i>df</i>
<i>First factor</i>				
Mod. for mutual dependence	-8015.72	36		
Mod. for $X(t, m_X) \perp\!\!\!\perp Y(t, m_Y) \mid Z$	-8505.97	18	980.50	18
Mod. for $Y(t, m_Y) \perp\!\!\!\perp X(t, m_X) \mid Z$	-8321.85	26	612.26	10
Mod. for $X(t, m_X) \perp\!\!\!\perp Y(t, m_Y) \mid Z$	-8095.38	31	159.32	5
Reduced model for mutual dependence	-8018.27	34	5.10	2
<i>Second factor</i>				
Full model	-997.32	0	0.00	0
Reduced model $Z_1 Z_3, Z_2 Z_3$	-998.78	2	2.92	2

**Fig. 3.** Resulting graph.

reduced version with mutual dependence between the two MPPs. The resulting graph is in figure 3, where the edge between Z_1 and Z_2 as well as the arrow from Z_2 to $X(t, m_X)$ are missing.

5. Concluding remarks

In this paper the class of model called graphical duration models as an extension of graphical models has been defined. This extension is not only purely theoretical, but it would permit the application of the language of graphical models in many longitudinal studies, for example to event history analysis and duration data; fields of applications are obviously that of duration models and event history analysis. The advantage in using

graphical duration models is to provide precise, parsimonious pictorial representations of the mutual relations between the random quantities, so that really complex model for marked point processes may be properly interpreted for a general public. Moreover, the graph suggests a convenient recursive factorization of the likelihood function, to pick out an efficient computational algorithms for maximization of the likelihood function.

Throughout the present paper, the continuous time context has been analyzed; however, the results are still valid for discrete time analysis; moreover, only a full parametric specification of the likelihood function has been treated here; in some simple cases, for example survival models, a nonparametric specification can be used, while a partial specification, such as in Cox models, requires a further assumption on the irrelevance of the unspecified part on conditional independence. Finally, in the specific of survival data, the graphical duration model that could be formulated by this approach is exactly conformable with the model achievable with classical graphical model using the general approach of transforming the hazard function as a linear function of explanatory variables.

As for classical graphical models, graphical duration models use is twofold: given a graph, to specify the collection of conditional independence and hence the joint probability distribution, and given a complex model, simplify the complete graph on the basis of inferential results for illustrating the relationships among random quantities. In the first use the structure of conditional independence is assumed known, while the second use applies model selection procedures in order to find suitable representations. The paper deals with both aspect, even if the applied examples focus on the latter one. Note that the asymmetry resulting from conditional local independence makes graphical duration models in some way similar to causal graph and suggests a causal interpretation of relationships which should be deepened indeed.

Appendix

Proof of proposition 1

For simplicity, the proof is given for a single point-marked process; extension to a more general case is immediate. Let \mathcal{M}_X^W be the subset of marks $m_{X,i}$ such that $i \in \mathcal{S}_Y$; let \mathcal{M}_Y^W be the subset of marks $m_{Y,k}$ such that $k \in \mathcal{S}_X$. Suppose that \mathcal{M}_X^W and \mathcal{M}_Y^W partition \mathcal{M}_W , such that $\mathcal{M}_X^W \cup \mathcal{M}_Y^W = \mathcal{M}_W$ and $\mathcal{M}_X^W \cap \mathcal{M}_Y^W = \emptyset$. If condition (3) is satisfied, then

$$\begin{aligned} f_W(t) &= \prod_{m_W \in \mathcal{M}_W} f_W(t, m_W)^{\xi_{m_W}} = \prod_{m_W \in \mathcal{M}_X^W} f_W(t, m_W) \prod_{m_W \in \mathcal{M}_Y^W} f_W(t, m_W) \\ &= \prod_{m_X \in \mathcal{M}_X} f_X(t, m_X) \prod_{m_Y \in \mathcal{M}_Y} f_Y(t, m_Y) = f_X(t) f_Y(t). \end{aligned}$$

where $f_W(\cdot)$ is the density function for the marked point process $W(t, m_W)$ constructed on the product-space of states of $X(t, m_X)$ and $Y(t, m_Y)$, whose density functions are

$f_X(\cdot)$ and $f_Y(\cdot)$ respectively. On the other hand, if condition (4) is satisfied, then

$$\begin{aligned} f_W(t) &= \prod_{m_W \in \mathcal{M}_W} f_W(t, m_W) = \prod_{m_W \in \mathcal{M}_X^W} f_W(t, m_W) \prod_{m_W \in \mathcal{M}_Y^W} f_W(t, m_W) \\ &= \prod_{m_W \in \mathcal{M}_X^W} f_W(t, m_W) \prod_{m_Y \in \mathcal{M}_Y} f_Y(t, m_Y) = f_{X|Y}(t) f_Y(t). \end{aligned}$$

where $f_{X|Y}(\cdot)$ is the conditional density function of $X(t, m_X)$ given $Y(t, m_Y)$. \square

References

- Aalen O.O. (1987) Dynamic modeling and causality. *Scandinavian Actuarial Journal*, 177-190.
- Aalen O.O., Borgan Ø., Keiding N. and Thormann J. (1980) Interaction between life history events. Non-parametric analysis for prospective and retrospective data in presence of censoring. *Scandinavian Journal of Statistics*, 7, 161–171.
- Arjas E. (1989) Survival models and martingale dynamics (with discussion) *Scandinavian Journal of Statistics*, 16, 177-225.
- Basu A. P. and Sun K. (1997) Multivariate exponential distributions with constant failure rates. *Journal of Multivariate Analysis*, 61, 159–169.
- Blossfeld, H.P., Klijzing E., Pohl K. and Rhwer G. (1995) Modelling Parallel Processes in Demography. An Application Example of the Causal Approach to Interdependent Systems. *3rd European Population Conference on Event History Analysis in Demography*, Milano.
- Cox D. R. (1972) Regression models and life-tables. *Journal of Royal Statistics Society*, 34, 187–220.
- Cox, D.R. and Lewis, P.A.W. (1972) Multivariate point processes. in *Proceedings of the sixth Berkeley Symposium on Mathematical Statistics and Probability*, vol. III.
- Cox D.R. and Wermuth N. (1996) *Multivariate Dependencies. Models, analysis and interpretation*. Chapman and Hall, London.
- Didelez V. (2001) Graphical Models for Event History Analysis based on Local Independence. Discussion Paper Nr. 158, University of Munich, Institute of Statistics.
- Edwards D. (2000) *Introduction to graphical modeling* Springer–Verlag.
- Gottard A. (1999) The analysis of Dynamic Interdependence between Fertility and Female Employment by Graphical Models, in *Proceedings of the 14th International Workshop on Statistical Modelling*, Graz, Austria.

- Lauritzen S.L. (1996) *Graphical Models*. Oxford: Oxford Science Publications.
- Lauritzen, S.L. and Wermuth, N. (1989) Graphical models for the association between variables, some of which are qualitative and some quantitative. *Annals of Statistics*, 17,31–57.
- Schweder, T. (1970) Composable Markov processes. *Journal of Applied Probability*, 7, 400–410.
- Tuma, N.B. and Hannan, M.T. (1984) *Social Dynamics, Models and Methods*. Orlando et al., Academic Press.
- van den Berg G.J. (1997) Association measures for durations in bivariate hazard rate models. *Journal of Econometrics*, 79, 221–245.
- Whittaker J.L. (1990) *Graphical Models in Applied Multivariate Statistics*. Wiley: New York.

Copyright © 2002

Anna Gottard