ELSEVIER

A multivariate counting process with Weibull-distributed first-arrival times

Francis Tuerlinckx*

Department of Psychology, University of Leuven, Tiensestraat 102, B-3000 Leuven, Belgium

Abstract

In this paper, we study a method to construct a multivariate counting process with positive dependencies between the event occurrences. Conditional on a random effect with a positive stable distribution, the univariate counting processes are independent non-homogeneous Poisson processes with a power intensity function. The applicability of the model is illustrated in three examples: a horse race model with several dependent channels, a dependent parallel-counter model and an interactive coactivation model.
© 2003 Elsevier Inc. All rights reserved.

Keywords: Response time models; Accumulator models; Coactivation model; Race models; Copula; Random effect; Poisson process

1. Introduction

Multivariate counting processes underly many response time models in psychology (for overviews, see Luce, 1986; Smith, 2001; Townsend & Ashby, 1983). For example, race (or parallel-counter) models, a class of successful models for choices and response times in speeded forced choice tasks, are based on such counting processes (Cousineau, 2004; Luce, 1986; Pike, 1966, 1973; Townsend & Ashby, 1983; Van Zandt, Colonius, & Proctor, 2000). In race models, discrete events are generated by several sources (or, alternatively, there are many channels; both terms are used interchangeably here) and the number of events originating from the different sources are tracked by separate counters. The counter that first reaches a prespecified criterion wins the race and determines the response. Commonly it is assumed that the events in one counter arrive independently from the events in another counter.

Another example of a response time model based on a multivariate counting process is the coactivation model to explain the redundant-targets effect in a detection task (Miller, 1982). Responses to redundant signals are typically faster than to non-redundant signals and this phenomenon can be modeled with a (channel summa-

tion) coactivation model in which the events from the different sources (each source corresponding with a particular target) are pooled and drive a single counter (Schwarz, 1989; Diederich, 1995); a response follows if the counter reaches a criterion. Again, it is typically assumed that the events from the different sources arrive independently.

In this paper, a multivariate counting process will be constructed in which the counts are positively correlated. Although the multivariate counting process in itself may be of interest in some applications, it will be considered mainly here as a useful vehicle to derive models for cognitive processes. The approach will be illustrated in this paper for the race and coactivation models that have been described in the preceding paragraphs.

For both examples (race and coactivation models), it is the case that assuming independence between the several counting processes simplifies the modeling task greatly. Thus, independence is often taken for granted for reasons of mathematical tractability, but it is unlikely that an independence assumption will always hold in reality. In the context of models for speeded forced choice, Usher and McClelland (2001) have recently presented a choice model in which the channels show lateral inhibition, thereby abolishing the independence assumption, and the latter authors motivate this mechanism by relying on neurophysiological evidence.

*Fax: +32-16-32-5916.
E-mail address: francis.tuerlinckx@psy.kuleuven.ac.be.

Huber and Cousineau (2004) also called for a race model with lateral inhibition to explain their results on short-term priming. Although lateral inhibition corresponds with negative dependency rather than with positive dependency, the approach presented in this paper yields positive dependencies, but it may be considered as a first step toward analytically tractable and fairly general multivariate counting process with an arbitrary degree of dependency.

Not only in the context of race models, but also for coactivation models, there is a need to allow for dependencies between the incoming events. For example, recent neurophysiological research (Giard & Peronnet, 1999; Fort, Delpuech, Pernier, & Giard, 2002) shows that for redundant multisensory signals, intersensory interactions occur relatively early during processing and that the results cannot be explained from an independent coactivation model.

It should be noted that for race models, there is an additional reason to assume independence, besides mathematical tractability: The dependency structure may not be identifiable. The problem is that if in a race process only the winner and the corresponding response time are observed, then one can always find an independence model that fits the data, even if in the true model the races are correlated. The identifiability problem has been studied by Cox (1959), Dzhafarov (1993), Marley and Colonius (1992), Townsend (1976), and Tsiatis (1975) and it has been an impediment to the active investigation of counter models with dependency between the event-generating sources. However, the fact that there are mathematical problems in identifying the dependence structure does not imply that the race processes are truly independent. Moreover, recent results in the statistical literature show that under some conditions (see below) dependent race models with counters containing only a single vacant slot (so-called horse race models) can be identified (Carrière, 1995; Escarela & Carrière, 2003). When presenting results on dependent race models, special attention will be given to this identifiability problem.

The remainder of the paper is organized as follows. First, we will introduce the *non-homogeneous Poisson process with a power intensity function* (power function NHPP) and some related results. Second, a specific method to induce dependencies between the first-arrival waiting time distributions, the copula method, is discussed. It is then shown how the specific type of copula method used in this paper can be framed as a random effects model. Subsequently, this random-effects approach is used in Section 3 to build a multivariate counting process in which the univariate marginal counting processes are conditional power function NHPPs. Fourth, the methods and techniques developed in the previous chapters are applied to three examples. The first example concerns a simple horse race

model between several dependent channels, the second one deals with a parallel-counter model with two dependent racing counters and the third one introduces a coactivation model with dependencies (i.e., an interactive coactivation model). The paper is closed with a general conclusion and discussion.

2. From a non-homogeneous Poisson process …

Consider a non-homogeneous Poisson process (NHPP) $\{N(t), t \geqslant 0\}$ with rate or intensity function $a(t)$ and $N(0) = 0$. The integrated rate $A(t)$ is defined as $\int_0^t a(s)\,ds$. It can be shown (Ross, 1996; Smith & Van Zandt, 2000) that the probability of observing exactly n events in a time interval of length t starting at time s equals

$$\Pr(N(s+t) - N(s) = n)$$
$$= \frac{e^{-(A(s+t) - A(s))}(A(s+t) - A(s))^n}{n!}. \tag{1}$$

Let the random variable T_1 denote the waiting time until the first event. The probability of observing a waiting time longer than t until the first event is equal to

$$\Pr(T_1 > t) = \Pr(N(t) = 0) = e^{-A(t)}. \tag{2}$$

Eq. (2) is the survivor function of the first-arrival waiting time $S_{T_1}(t)$. In general, the survivor function of the waiting time until the nth event, denoted by the random variable T_n, is as follows:

$$S_{T_n}(t) = \Pr(T_n > t) = \Pr(N(t) < n)$$
$$= \sum_{j=0}^{n-1} \frac{e^{-A(t)} A(t)^j}{j!} = \frac{\Gamma(n, A(t))}{\Gamma(n)}, \tag{3}$$

where $\Gamma(n, x) = \int_x^\infty u^{n-1} e^{-u}\,du$ is the (complementary) incomplete gamma function. The equation on the right follows from using the expression for the finite sum $\sum_{j=0}^{n-1} \frac{x^j}{j!} = \frac{e^x \Gamma(n,x)}{\Gamma(n)}$ (Prudnikov, Brychkov, & Marichev, 1986). In a NHPP, the density function of the waiting time until the nth arrival reads as

$$f_{T_n}(t) = -\frac{dS_{T_n}(t)}{dt} = \frac{A(t)^{n-1} A'(t) e^{-A(t)}}{(n-1)!}. \tag{4}$$

Eq. (4) can be found from Eq. (3) by directly differentiating the terms of the finite sum or by applying Leibnitz's rule for differentiating an integral to the complementary incomplete gamma function (Spiegel, 1983). The density of the waiting time for the nth event is called a *non-homogeneous* or *time-varying gamma distribution* (Smith & Van Zandt, 2000).

If the event rate $a(t)$ is a power function of the form $\lambda \gamma t^{\gamma-1}$ (and thus $A(t) = \lambda t^\gamma$), then Eq. (2) is the survivor function of a (two-parameter) Weibull distribution:

$$S_{T_1}(t) = e^{-\lambda t^\gamma}.$$

The counting process reduces to a homogeneous Poisson process if $\gamma = 1$ and in that case Eq. (1) simplifies to a Poisson probability distribution with mean λt. For a homogeneous Poisson process, T_1, the waiting time until the first event, but also subsequent interarrival waiting times, are exponentially distributed with mean $\frac{1}{\lambda}$. Consequently, the waiting time until the nth event (see Eq. (4)) is gamma distributed with parameters n and λ. In the remainder of the paper, we will deal with the more general case $\gamma \neq 1$. In Appendix A, it is indicated how to simulate random variates from a power function NHPP.

In most applications, the use of the Weibull distribution is justified on asymptotical grounds because it is the limiting distribution of the minimum of a large number of parallel racing processes (for the latter approach, see Colonius, 1995; Cousineau, Goodman, & Shiffrin, 2002; Gumbel, 1958). The distributions of the racing processes should satisfy some conditions to guarantee convergence to the Weibull distribution (see Cousineau et al., 2002). Contrary to this asymptotic justification, a power function NHPP provides a finite-sample justification of the common two-parameter Weibull distribution because the latter can be derived as the first-arrival waiting time distribution from such an NHPP counting process. As a consequence, the shape of the Weibull distribution (determined by the parameter γ) is related to the rate of increase of the intensity function of the underlying counting process.

The model presented so far is a univariate counting process: There is only a single type of events and an associated event rate. To arrive at a multivariate counting process, more than one event-generating source has to be introduced (or, equivalently, more than one channel). Suppose that each source i ($i = 1, \ldots, I$) generates events according to a power intensity function $a_{(i)}(t) = \lambda_{(i)} \gamma_{(i)} t^{\gamma_{(i)} - 1}$. In such a case, there are I different event types and each arriving event carries its own label corresponding to the source by which it has been generated. It will be assumed that all rate functions are proportional to each other, such that $\gamma_{(1)} = \cdots = \gamma_{(I)} = \gamma$. For the time being, the I processes are considered to be independent.

Several useful properties can be derived from the independent multivariate counting process defined above. For example, starting from the I counting processes $\{N_{(i)}(t), t \geq 0\}$, we can define a new counting process $\{N(t), t \geq 0\}$ that is the superposition of the I individual counting processes: $N(t) = N_{(1)}(t) + \cdots + N_{(I)}(t)$. Concerning this new counting process, the following proposition can be proven.

Proposition 1. *The counting process $\{N(t), t \geq 0\}$ defined as the superposition of I independent power function NHPPs $\{N_{(i)}(t), t \geq 0\}$, with rate functions $a_{(i)}(t) = \lambda_{(i)} \gamma t^{\gamma - 1}$, is again a power function Poisson process, and its rate function is the sum of the separate intensity functions, $a(t) = \sum_i^I a_{(i)}(t) = (\lambda_{(1)} + \cdots + \lambda_{(I)}) \gamma t^{\gamma - 1}$.*

Proof. This proposition is proven easily using the Laplace transform for a power function NHPP and the independence of the I processes. $\square$

Because the new process $N(t)$ is again a power function NHPP, the waiting time until the first arriving event (irrespective of the label it is carrying) is Weibull distributed with parameters $\lambda_{(1)} + \cdots + \lambda_{(I)}$ and γ. The same result is obtained from the fact that the minimum of I Weibull distributed random variables with an equal power parameter γ but possibly a different $\lambda_{(i)}$ is also Weibull distributed with parameters $\lambda_{(1)} + \cdots + \lambda_{(I)}$ and γ; this is the so-called *min-stable property* of the Weibull distribution. In addition, if we do not ignore the labels of the arriving events, we can find an expression for the probabilities that a particular source i generated the first arriving event, as is done in Proposition 2.

Proposition 2. *Given I independent power function NHPPs $\{N_{(i)}(t), t \geq 0\}$, the probability that source i has generated the first arriving event equals $\frac{\lambda_{(i)}}{\lambda_{(1)} + \cdots + \lambda_{(I)}}$. The waiting time for this first event is Weibull distributed with parameters $\lambda_{(1)} + \cdots + \lambda_{(I)}$ and γ.*

Proof. First, we verify the probability formula, making use of the fact that the I Poisson processes are independent:

$$
\begin{aligned}
&\Pr(\text{event of type } i \text{ arrives first}) \\
&= \Pr(T_{1(i)} < T_{1(j)}, j \neq i) \\
&= \int_0^\infty f_{T_{1(i)}}(t_i) \prod_{j \neq i}^I \left(\int_{t_i}^\infty f_{T_{1(j)}}(t_j)\, dt_j \right) dt_i \\
&= \int_0^\infty \lambda_{(i)} \gamma t_i^{\gamma-1} \exp(-\lambda_{(i)} t_i^\gamma) \\
&\quad \times \prod_{j \neq i}^I \exp(-\lambda_{(j)} t_i^\gamma)\, dt_i \\
&= \frac{\lambda_{(i)}}{\lambda_{(1)} + \cdots + \lambda_{(I)}} \int_0^\infty \left(\sum_{j=1}^I \lambda_{(j)} \right) \\
&\quad \times \gamma t_i^{\gamma-1} \exp\left(-\left(\sum_{j=1}^I \lambda_{(j)} \right) t_i^\gamma \right) dt_i \\
&= \frac{\lambda_{(i)}}{\lambda_{(1)} + \cdots + \lambda_{(I)}},
\end{aligned}
\tag{5}
$$

where $T_{1(i)}$ and t_i denote the random variable and its realization, respectively, of the first-arrival waiting time for process i. Next, we derive the distribution

for the minimum of the I first-arrival waiting times. If $T = \min_i(T_{1(i)})$, then the survivor function of T equals

$$
\begin{aligned}
\Pr(T > t) &= \Pr(\min_i(T_{1(i)}) > t) \\
&= \Pr(T_{1(1)} > t, \ldots, T_{1(I)} > t) \\
&= \prod_{i=1}^{I} \Pr(T_{1(i)} > t) \\
&= \exp(-(\lambda_{(1)} + \cdots + \lambda_{(I)})t^{\gamma}).
\end{aligned}
$$

This survivor function is called the *overall survivor function* and it is equal to the Weibull distribution given in the proposition. $\square$

The model from Proposition 2 is a horse race model with I several independent channels. Stated otherwise, there are I counters performing a race and each counter has only a single free slot so that the first incoming event determines the response. A special property of this horse race model is that the label and time of the first-arrival are independent.

The usefulness of a Poisson process with a power intensity function can be questioned. The power intensity function is adopted here mainly because it is more general than its more commonly used special case, the constant intensity function leading to the regular homogeneous Poisson process. An additional advantage of the power intensity function is that it has a straightforward relation to the common Weibull distribution. However, the results presented in this paper can be generalized easily to a more general NHPP in which the integrated intensity functions are of the following form: $A_{(i)}(t) = \lambda_{(i)}h(t)$, where $h(t)$ is a known standard integrated intensity function. All previously presented and subsequent results still hold if in the equations t^{γ} and $\gamma t^{\gamma-1}$ are replaced by $h(t)$ and $h'(t)$, respectively. Because there are no restrictions for $h(t)$, it can be any function. An interesting specification for $h(t)$ is a non-monotonic intensity function for which there is a baseline intensity that increases and then returns again to baseline. Such a function may be used to match the time-varying intensity typically inferred from neural spike train data (see e.g., Ritov, Raz, & Bergman, 2002; Sanger, 2002). It should be noted however that when t^{γ} is replaced by $h(t)$, the first-arrival waiting time is not necessarily Weibull distributed anymore.

In the rest of the paper, we will relax the assumption of independence among the I counting processes. In order to arrive at a plausible model for dependencies among the several counting processes, we will make a detour via copulas which are discussed in the next section.

3. … to copulas …

In this section, we will construct a joint distribution for the I-dependent first-arrival waiting times $T_{1(i)}$ ($i = 1, \ldots, I$) such that each univariate marginal distribution of a first-arrival waiting time $T_{1(i)}$ is still Weibull distributed with parameters $\lambda_{(i)}$ and γ. To achieve this goal, we use copulas (Genest and MacKay, 1986; Joe, 1997). As suggested already by its name, a copula is a function that "couples" several univariate survivor functions to form a multivariate survivor function. Thus, the joint survivor function of $T_{1(1)}, \ldots, T_{1(I)}$, denoted as $S_{T_{1(1)}, \ldots, T_{1(I)}}(t_1, \ldots, t_I)$, is defined on the basis of the separate univariate marginal survivor functions $S_{T_{1(i)}}(t_i)$'s and some copula function C as follows:

$$
S_{T_{1(1)}, \ldots, T_{1(I)}}(t_1, \ldots, t_I) = C(S_{T_{1(1)}}(t_1), \ldots, S_{T_{1(I)}}(t_I)), \quad (6)
$$

where the copula function has to satisfy certain requirements (see Definition 1). The copula function C captures the dependency between the I random variables. A copula is a convenient way of modeling dependencies between random variables because it separates the marginal and the association part of the multivariate distribution. Choosing a different copula function will lead to a different joint survivor function. However, for a given joint survivor function (with continuous margins), one can always find a unique copula representation; this is known as Sklar's theorem (Jouini & Clemen, 1996).

In this paper, the arguments of the copula function are survivor functions and a multivariate survivor function follows. However, the copula is actually a multivariate *cumulative distribution* function (CDF) and Definition 1 contains a formal definition of the copula function.

Definition 1. An I-dimensional copula C is a multivariate CDF with support of the I-dimensional hypercube $[0, 1]^I$. The univariate marginal CDFs of the copula are uniform on $[0, 1]$.

Because of its definition as a multivariate CDF, the copula has all properties of a usual multivariate CDF. Thus, it has to hold that $C(u_1, \ldots, u_{i-1}, 0, u_{i+1}, \ldots, u_I) = 0$ (with $0 \leqslant u_j \leqslant 1$). Moreover, since the univariate marginal CDFs are uniform on $[0, 1]$, it is the case that $C(1, \ldots, 1, u_i, 1, \ldots, 1) = u_i$.

We will now show how a multivariate survivor function can be constructed given I univariate marginal survivor functions $S_{T_{1(i)}}(t_i)$ and given a copula C.

Proposition 3. *The function $C(S_{T_{1(1)}}(t_1), \ldots, S_{T_{1(I)}}(t_I))$ is a multivariate survivor function for the random variables $T_{1(1)}, \ldots, T_{1(I)}$ provided $S_{T_{1(i)}}(t_i)$ has a unique inverse ($i = 1, \ldots, I$).*

Proof. If $T_{1(i)}$ is a random variable with survivor function $S_{T_{1(i)}}(t_i)$, then it holds that $U_i = S_{T_{1(i)}}(T_{1(i)})$ is uniformly distributed on $[0,1]$ (this is the probability integral transform) and $u_i = S_{T_{1(i)}}(t_i)$ is a realization from this uniform distribution. Because the copula is a multivariate CDF, it can be written as

$$C(u_1, \ldots, u_I) = \Pr(U_1 \leqslant u_1, \ldots, U_I \leqslant u_I)$$

with U_i uniformly distributed on $[0,1]$. Then, it can be seen that

$$\begin{aligned}
&C(S_{T_{1(1)}}(t_1), \ldots, S_{T_{1(I)}}(t_I)) \\
&= \Pr(U_1 \leqslant S_{T_{1(1)}}(t_1), \ldots, U_I \leqslant S_{T_{1(I)}}(t_I)) \\
&= \Pr(S_{T_{1(1)}}^{-1}(U_1) \geqslant t_1, \ldots, S_{T_{1(I)}}^{-1}(U_I) \geqslant t_I) \\
&= \Pr(T_{1(1)} \geqslant t_1, \ldots, T_{1(I)} \geqslant t_I) \\
&= S_{T_{1(1)}, \ldots, T_{1(I)}}(t_1, \ldots, t_I),
\end{aligned}$$

where the inequality signs are reversed when going from the first to the second line because $S_{T_{1(i)}}^{-1}$ is a monotonic decreasing function. $\quad\square$

A simple example of a copula function is the independence copula: $C(u_1, \ldots, u_I) = \prod_{i=1}^{I} u_i$. If the independence copula is applied to univariate marginal survivor functions, one obtains a multivariate survivor function with independence between components. In this paper we deal with the following copula: $C(u_1, \ldots, u_I) = e^{-s^\alpha}$ and $s = \sum_{i=1}^{I}(-\log(u_i))^{\frac{1}{\alpha}}$, with $0 < \alpha \leqslant 1$. This copula is called the Gumbel–Hougaard copula (Gumbel, 1961; Hougaard, 1986a, b). It can be derived that $C(u_1, \ldots, u_I) = \phi(\sum_{i=1}^{I}\phi^{-1}(u_i))$, with $\phi(s) = e^{-s^\alpha}$ and $\phi^{-1}(u) = (-\log(u))^{\frac{1}{\alpha}}$.

The joint survivor function defined through the Gumbel–Hougaard copula with I univariate marginal Weibull survivor functions as arguments gives a multivariate Weibull distribution:

$$\begin{aligned}
&S_{T_{1(1)}, \ldots, T_{1(I)}}(t_1, \ldots, t_I) \\
&= \phi(\phi^{-1}(S_{T_{1(1)}}(t_1)) + \cdots + \phi^{-1}(S_{T_{1(I)}}(t_I))) \\
&= \exp\left\{ -\left(\lambda_{(1)}^{\frac{1}{\alpha}} t_1^{\frac{\gamma}{\alpha}} + \cdots + \lambda_{(I)}^{\frac{1}{\alpha}} t_I^{\frac{\gamma}{\alpha}} \right)^\alpha \right\} \\
&\quad (0 < \alpha \leqslant 1).
\end{aligned} \tag{7}$$

The condition $0 < \alpha \leqslant 1$ has to be imposed in order to obtain a proper multivariate probability density.

The parameter α captures the dependency between the I random variables. If $\alpha = 1$, the independence copula follows as a special case of the Gumbel–Hougaard copula because then the joint survivor function simplifies to a product of the I univariate marginal Weibull survivor functions. On the other hand, if $\alpha \to 0$, the joint survivor function becomes $\exp\{-\max_i(\lambda_{(i)} t_{1(i)}^\gamma)\} = \min_i(S_{T_{1(i)}}(t_{1(i)}))$, which corresponds to the complete positive dependence case. Thus, the multivariate Weibull

model constructed through the Gumbel–Hougaard copula is restricted to have positive dependence between the random variables.

Let us assume for the moment that the function $\phi(s) = e^{-s^\alpha}$ can be considered as a Laplace transform of some unknown probability density $g(\theta)$, that is $\phi(s) = E(e^{-s\Theta}) = \int_0^\infty e^{-s\theta}\, dG(\theta)$. In that case, the Gumbel–Hougaard copula we use can be rewritten as follows (see Marshall & Olkin, 1988; Oakes, 1989):

$$\begin{aligned}
&\phi\left(\sum_{i=1}^{I} \phi^{-1}(S_{T_{1(i)}}(t_i)) \right) \\
&= \int_0^\infty \exp\left\{ -\theta\left(\sum_{i=1}^{I} \phi^{-1}(S_{T_{1(i)}}(t_i)) \right) \right\} dG(\theta) \\
&= \int_0^\infty \exp\left\{ -\theta\left(\sum_{i=1}^{I} \lambda_{(i)}^{\frac{1}{\alpha}} t_i^{\frac{\gamma}{\alpha}} \right) \right\} dG(\theta) \\
&= \int_0^\infty \prod_{i=1}^{I} W_{T_{1(i)}}(t_i)\, dG(\theta)
\end{aligned} \tag{8}$$

with $W_{T_{1(i)}}(t_i)$ a Weibull survivor function with parameters $\theta \lambda_{(i)}^{\frac{1}{\alpha}}$ and $\frac{\gamma}{\alpha}$. Note that this implies that θ has to be positive and thus that $G(\theta)$ must be CDF of a positive random variable.

From Eq. (8), one can derive that if $\phi(s)$ is indeed the Laplace transform of $g(\theta)$, then the dependency structure imposed on the random variables $T_{1(i)}$ ($i = 1, \ldots, I$) through the copula can actually be understood in terms of a random effect with density $g(\theta)$. Conditionally on the random effect θ, the I Weibull random variables are independent; this is known as the conditional independence assumption. The random effect θ is also called a *frailty* (Oakes, 1989), a term used in failure time analysis because if the first-arriving event represents a failure, θ can be conceived of as a latent predisposition to fail. The random effect or frailty is an unobserved component of the intensity function which is assumed to be drawn randomly from a population distribution. The fact that the same unobserved component is shared by the intensity functions of the several counting process induces the dependency.

We have assumed up until now that there exists a $g(\theta)$ which has $\phi(s) = e^{-s^\alpha}$ as its Laplace transform, but is this true? It turns out that $\phi(s) = e^{-s^\alpha}$ is the Laplace transform of a *positive stable density* (Feller, 1971, p. 165):

$$g(\theta) = \begin{cases} \dfrac{1}{\pi}\sum_{k=1}^{\infty} \dfrac{\Gamma(k\alpha+1)}{k!}(-1)^{k+1}\theta^{-\alpha k-1}\sin(\alpha k\pi) \\ \quad \text{if } \alpha < 1, \\ \delta_{\theta,1} \\ \quad \text{if } \alpha = 1, \end{cases} \tag{9}$$
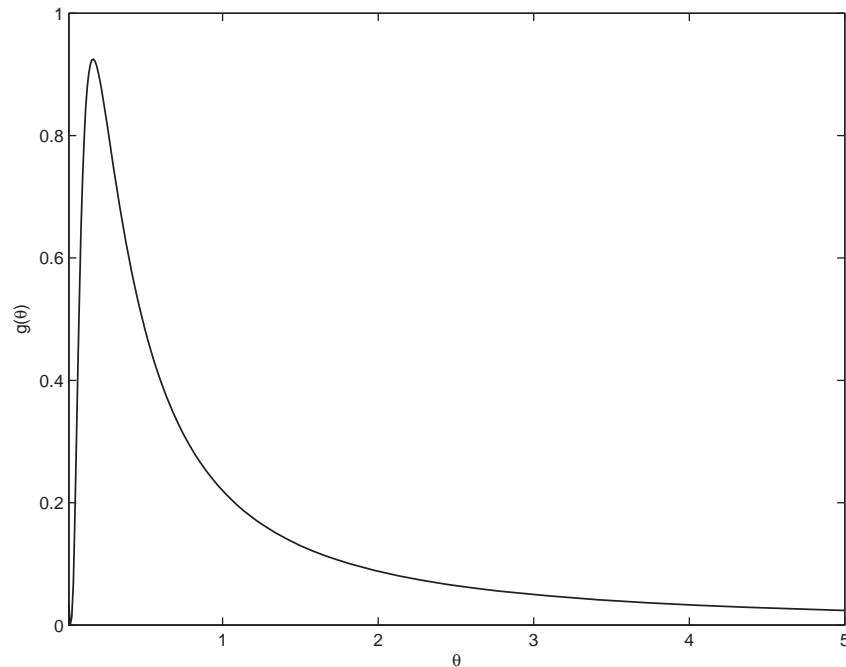
Fig. 1. The density function for a positive stable distribution with $\alpha = 0.5$.

where $\delta_{\theta,1}$ is the Kronecker delta taking a value of 1 if $\theta = 1$ and zero otherwise. A positive stable density is defined only for positive values of θ, and is characterized by a positive skewness, very heavy tails and the fact that only moments $E(\Theta^q)$ for $q < \alpha$ exist (see Fig. 1 for an example). Because of the restriction we have placed on α, the distribution has no mean or variance. The infinite sum in Eq. (9) converges fast for large values of θ, but it may be slow for values of θ close to zero. Appendix B contains some additional information on the positive stable density and in Appendix C a simple algorithm is described for generating random variates from a positive stable distribution.

4. ... to a conditional non-homogeneous Poisson process

In the previous section, we have explained how a multivariate first-arrival waiting times distribution can be constructed from the univariate marginal waiting time distributions, which were all Weibull distributions, and from a specific dependency structure, called the Gumbel–Hougaard copula. Because the function $\phi(s) = e^{-s^\alpha}$ is the Laplace transform of a known probability distribution, the multivariate Weibull model could be casted as a random-effects model. In a next step, we will make use of this random-effects interpretation to induce dependencies between several Poisson processes. Because a Weibull process is the first-arrival waiting time distribution in a power function NHPP, we also have the possibility to construct, conditionally on the frailty, several independent power function NHPPs.

After integrating out the random effect, the several NHPPs become dependent, and this leads to a multivariate counting process. The counting process is referred to as a *multivariate conditional power function NHPP*.

From Eq. (8), conditional on the frailty θ, we know that in the multivariate Weibull distribution, the first-arrival waiting times are independently Weibull distributed. We also know that the Weibull distribution follows from a Poisson model with a power intensity function. Therefore, let us consider a Poisson process with integrated intensity function $A_{(i)}(t) = \theta \lambda_{(i)}^{\frac{1}{\alpha}} t^{\frac{\gamma}{\alpha}}$. For this Poisson process, the first-arrival times are Weibull distributed with parameters $\theta \lambda_{(i)}^{\frac{1}{\alpha}}$ and $\frac{\gamma}{\alpha}$. Conditional on the frailty, the probability of having observed n events of type i at time t equals (see Eq. (1))

$$
\Pr(N_{(i)}(t) = n | \theta)
$$
$$
= \frac{\exp\left(-\theta \lambda_{(i)}^{\frac{1}{\alpha}} t^{\frac{\gamma}{\alpha}}\right) \left(\theta \lambda_{(i)}^{\frac{1}{\alpha}} t^{\frac{\gamma}{\alpha}}\right)^n}{n!}. \tag{10}
$$

Marginally the first-arrival waiting times are still Weibull distributed (because of the definition of the copula), but the marginal counting process is not a Poisson process anymore (therefore, it is called a conditional Poisson process), unless in the trivial case that $\alpha = 1$. This can be seen by deriving the marginal probability (i.e., with θ integrated out) for the event $N_{(i)}(t) = n$.

Proposition 4. *For a multivariate conditional power function NHPP as defined by Eq. (10) and θ having a positive stable distribution with parameter α, the marginal probability $\Pr(N_{(i)}(t) = n)$ equals*

$$\Pr(N_{(i)}(t) = n)$$

$$= (-1)^n \frac{\left(\lambda_{(i)}^{\frac{1}{\alpha}} t^{\frac{\gamma}{\alpha}}\right)^n}{n!} \left\{\frac{d^n e^{-s^\alpha}}{ds^n}\bigg|_{s = \lambda_{(i)}^{\frac{1}{\alpha}} t^{\frac{\gamma}{\alpha}}}\right\}.$$

Proof. To obtain the marginal probability, we have to integrate $\Pr(N_{(i)}(t) = n|\theta)$ over the positive stable distribution of θ with parameter α:

$$\Pr(N_{(i)}(t) = n) = \int_0^\infty \Pr(N_{(i)}(t) = n|\theta)g(\theta)\,d\theta$$

$$= \frac{\left(\lambda_{(i)}^{\frac{1}{\alpha}} t^{\frac{\gamma}{\alpha}}\right)^n}{n!} \int_0^\infty \theta^n$$

$$\times \exp\left(-\theta \lambda_{(i)}^{\frac{1}{\alpha}} t^{\frac{\gamma}{\alpha}}\right) g(\theta)\,d\theta.$$

An explicit solution to the integral can be obtained by using a special theorem on Laplace transforms (Spiegel, 1983). If the Laplace transform of $g(\theta)$ is $\phi(s)$, then the Laplace transform of $\theta^n g(\theta)$ equals $(-1)^n \frac{d^n \phi(s)}{ds^n}$. Applying this theorem in our case gives the desired result. □

It is easy to see that the probability of observing no event until time t equals $e^{-\lambda_{(i)} t^\gamma}$, which was already apparent from the fact that the marginal first-arrival times are Weibull distributed. One can check that the marginal probabilities of observing one or more events do not equal anymore the probabilities that follow from an NHPP.

The unconditional joint probability is derived in Proposition 5.

Proposition 5. *For a multivariate conditional power function NHPP as defined by Eq. (10) and θ having a positive stable distribution with parameter α, the unconditional joint probability $\Pr(N_{(1)}(t_1) = n_1, \ldots, N_{(i)}(t_i) = n_i, \ldots, N_{(I)}(t_I) = n_I)$ (with $\sum_{i=1}^I n_i = n$) is as follows:*

$$\left(\prod_{i=1}^I \frac{\left(\lambda_{(i)}^{\frac{1}{\alpha}} t_i^{\frac{\gamma}{\alpha}}\right)^{n_i}}{n_i!}\right)(-1)^n \left\{\frac{d^n e^{-s^\alpha}}{ds^n}\bigg|_{s = \sum_{i=1}^I \lambda_{(i)}^{\frac{1}{\alpha}} t_i^{\frac{\gamma}{\alpha}}}\right\}.$$

Proof. The proof is analogous to the derivation of the marginal probability with the only difference being that the conditional joint probability is considered as the

starting point:

$$\Pr(N_{(1)}(t_1) = n_1, \ldots, N_{(i)}(t_i) = n_i, \ldots, N_{(I)}(t_I) = n_I|\theta)$$

$$= \prod_{i=1}^I \Pr(N_{(i)}(t) = n|\theta). \quad \square$$

In the previous and in subsequent equations, the nth derivative of $\phi(s) = e^{-s^\alpha}$ has to be calculated. This can be done easily using the recurrence relation given in Proposition 6.

Proposition 6. *If we define the nth derivative of e^{-s^α} to be $D_n = \frac{d^n e^{-s^\alpha}}{ds^n}$ and $D_0 = e^{-s^\alpha}$, then the following holds:*

$$D_n = -\alpha \sum_{i=0}^{n-1} \binom{n-1}{i}(\alpha - 1)$$

$$\cdots(\alpha - (n-1) + i)s^{\alpha - n + i} D_i,$$

where $(\alpha - 1)\cdots(\alpha - (n-1) + i) \equiv 1$ for $i = n - 1$.

Proof. We will give a proof by induction. For $n = 1$, $D_1 = -\alpha s^{\alpha - 1} e^{-s^\alpha} = -\alpha s^{\alpha - 1} D_0$, which proves the proposition for this case. Now suppose the proposition holds for n and we want to find D_{n+1}:

$$D_{n+1} = \frac{dD_n}{ds} = -\alpha \sum_{i=0}^{n-1} \binom{n-1}{i}$$

$$\times (\alpha - 1)\cdots(\alpha - (n-1) + i)$$

$$\times [(\alpha - n + i)s^{\alpha - n + i - 1} D_i + s^{\alpha - n + i} D_{i+1}]$$

$$= -\alpha \sum_{i=0}^{n-1} \binom{n-1}{i}(\alpha - 1)\cdots(\alpha - (n-1) + i)$$

$$\times (\alpha - n + i)s^{\alpha - n + i - 1} D_i - \alpha \sum_{i=0}^{n-1} \binom{n-1}{i}$$

$$\times (\alpha - 1)\cdots(\alpha - (n-1) + i)s^{\alpha - n + i} D_{i+1}$$

$$= -\alpha \sum_{i=0}^{n-1} \binom{n-1}{i}$$

$$\times (\alpha - 1)\cdots(\alpha - n + i)s^{\alpha - (n+1) + i} D_i$$

$$- \alpha \sum_{i=1}^{n} \binom{n-1}{i-1}$$

$$\times (\alpha - 1)\cdots(\alpha - (n-1) + i - 1)s^{\alpha - n + i - 1} D_i$$

$$= -\alpha(\alpha - 1)\cdots(\alpha - n)s^{\alpha - (n+1)} D_0$$

$$- \alpha \sum_{i=1}^{n-1} \left[\binom{n-1}{i} + \binom{n-1}{i-1}\right]$$

$$\times (\alpha - 1)\cdots(\alpha - n + i)s^{\alpha - n + i - 1} D_i - \alpha s^{\alpha - 1} D_n$$

$$= -\alpha \sum_{i=0}^{(n+1)-1} \binom{(n+1)-1}{i}$$

$$\times (\alpha - 1)\cdots(\alpha - ((n+1)-1) + i)s^{\alpha - (n+1) + i} D_i,$$

where we have made use of the equality $\binom{n-1}{i} + \binom{n-1}{i-1} = \binom{n}{i}$ for binomial coefficients (Abramowitz & Stegun, 1974, p. 822). This proves the proposition. $\square$

Because of this closed-form expression for the nth derivative, all computations concerning the multivariate counting process are relatively simple to carry out and there is no need to rely on approximation techniques such as numerical integration. Simulations from the multivariate conditional power function NHPP can be obtained by drawing first θ from a positive stable density (see Appendix C) and then simulating several independent power function NHPPs (see Appendix A) with the appropriate intensity function.

5. Three applications

In this section, we will illustrate the applicability of the multivariate counting process with positive dependencies in three examples. The first example is an extension of the independent horse race model from Proposition 2. In the second application, we allow for dependencies in the two-counter model of Smith and Van Zandt (2000). Third, we present a channel summation coactivation model (Diederich, 1995; Schwarz, 1989) with positive dependencies.

In all three examples, we construct a model for the decision time (and choice probabilities, if relevant) based upon the multivariate counting process. However, besides the decision time, it is commonly assumed (Luce, 1986) that the observed response time consists furthermore of a residual time component, which represents all nondecision processes (e.g., the time needed for encoding the stimulus and planning and executing the response). The residual response time is not considered explicitly in this paper, but the model equations need only a minor change in order to include a single residual time parameter (i.e., by replacing all occurrences of t by $t - T_{er}$, where T_{er} is the residual time component).

5.1. Application 1: a horse race between multiple dependent channels

Consider a horse race performed between I channels with positive dependency so that the first-arrival times follow a multivariate Weibull distribution. To derive the distribution of first-arrival waiting times and the choice probabilities, we could use the general result of Marley (1989), because his result also encompasses the multivariate Weibull distributions as a special case. However, here we present an alternative way of deriving the waiting time distribution and choice probability by first conditioning on the frailty θ and then integrating over $g(\theta)$.

Conditional on θ, we are in the case of I independent Poisson processes and we may use the results from Proposition 2. Thus, conditional on θ, the waiting time

until the first arrival is Weibull distributed with parameters $\theta(\sum_{i=1}^{I} \lambda_{(i)}^{\frac{1}{\alpha}})$ and $\frac{\gamma}{\alpha}$. Integrating over the distribution of θ gives the *overall* survivor function

$$S_T(t) = \int_0^\infty \exp\left\{ -\theta\left(\sum_{i=1}^{I} \lambda_{(i)}^{\frac{1}{\alpha}}\right) t^{\frac{\gamma}{\alpha}} \right\} dG(\theta)$$

$$= \exp\left\{ -\left(\sum_{i=1}^{I} \lambda_{(i)}^{\frac{1}{\alpha}}\right)^{\alpha} t^{\gamma} \right\}, \qquad (11)$$

which is again the survivor function of a Weibull distribution. In the derivation, we have applied the definition of the Laplace transform of a positive stable distribution. Thus, it has been shown that the multivariate Weibull distribution constructed with the Gumbel–Hougaard copula is min-stable because the minimum of several (dependent) Weibull distributions is still Weibull distributed. The choice probability can be found using the same logic. Conditional on θ, the choice probability can be derived from Eq. (5):

$$\Pr(W_{1(i)} < W_{1(j)}, j \neq i) = \frac{\theta \lambda_{(i)}^{\frac{1}{\alpha}}}{\theta \lambda_{(1)}^{\frac{1}{\alpha}} + \cdots + \theta \lambda_{(I)}^{\frac{1}{\alpha}}}$$

$$= \frac{\lambda_{(i)}^{\frac{1}{\alpha}}}{\lambda_{(1)}^{\frac{1}{\alpha}} + \cdots + \lambda_{(I)}^{\frac{1}{\alpha}}}, \qquad (12)$$

making use of the random variables $W_{1(i)}$ appearing in Eq. (8). Because the conditional choice probability is independent of θ, Eq. (12) is also the marginal probability.

As can be seen from Eqs. (11) and (12), the choice probability and choice time are independent, despite the dependence among the race processes. This follows also from the general result proven by Kochar and Proschan (1991) that the choice probability and the first-arrival time in a dependent channel model are independent if and only if the hazard rates[1] of the individual first-arrival time distributions are proportional. In our model, the hazard rate of the marginal distribution of the ith first-arrival times is equal to $\gamma \lambda_{(i)} t^{\gamma-1}$ and consequently the I hazard rates are clearly proportional to each other.

To illustrate the effect of the dependency parameter, let $\lambda_{(m)} = \max_i(\lambda_{(i)})$ and suppose this maximum is unique $(\lambda_{(m)} > \lambda_{(i)}, i \neq m)$. From Eq. (11) it can be deduced that if $\alpha \to 0$, the survivor function of the response time converges to $\exp\{-\max_i(\lambda_{(i)}) t^{\gamma}\} = \exp(-\lambda_{(m)} t^{\gamma})$. This implies that the overall survivor function converges to the first-arrival waiting time

---

[1] The hazard rate $b(t)$ is defined as $-\frac{d \log S_T(t)}{dt}$, where $S_T(t)$ is the survivor function (Luce, 1986). Hence for a Weibull distribution, the hazard rate is equal to the intensity rate of the associated power function NHPP.

survivor function of the source with the largest intensity function (i.e., the fastest process). Moreover, if $\alpha \to 0$, the probability of choosing the response option with the largest $\lambda_{(i)}$ goes to one. Then it can be deduced from Eq. (12) that

$$\lim_{\alpha \to 0} \Pr(W_{1(m)} < W_{1(i)}, i \neq m)$$

$$= \lim_{\alpha \to 0} \frac{\lambda_{(m)}^{\frac{1}{\alpha}}}{\lambda_{(1)}^{\frac{1}{\alpha}} + \cdots + \lambda_{(I)}^{\frac{1}{\alpha}}}$$

$$= \lim_{\alpha \to 0} \frac{1}{\left(\frac{\lambda_{(1)}}{\lambda_{(m)}}\right)^{\frac{1}{\alpha}} + \cdots + \left(\frac{\lambda_{(m)}}{\lambda_{(m)}}\right)^{\frac{1}{\alpha}} + \cdots + \left(\frac{\lambda_{(I)}}{\lambda_{(m)}}\right)^{\frac{1}{\alpha}}} = 1,$$

because $\frac{\lambda_{(i)}}{\lambda_{(m)}} \leqslant 1$, with equality if $i = m$. In the case of perfect dependency, it is the source with the largest rate function that determines the response. Stated otherwise, positive dependency among the counters suppresses the randomness in the choice response, but not in the response times. If $\alpha = 1$, the independent horse race model presented in Proposition 2 follows. In the case the maximum $\lambda_{(m)}$ equals one or more other λ parameters, the limiting probability becomes $\frac{1}{k+1}$, where k stands for the number of λ's equal to $\lambda_{(m)}$.

As mentioned before, an important issue in dependent race models is the identifiability of the model. A very important distinction in this respect is the one between parameter identifiability and process identifiability. From Eqs. (11) and (12), it can be deduced that all parameters (the $\lambda_{(i)}$'s and α) are identified (i.e., different values for the parameters lead to a different joint probability function); hence, there is parameter identification and it is possible to estimate all parameters uniquely from the data.

However, is there also identification on the level of the process? If there is process identification, the parameter α can be seen as a quantification of the amount of dependency between the I processes. Process identification is only possible if a single multivariate survivor function, as the one in Eq. (7), leads to the derived overall survivor function and the response probabilities. A counterexample to this conjecture is easily found by defining a horse race model with I independent channels and with $\lambda_{(i)}^{\frac{1}{\alpha}}(\sum_{j=1}^{I} \lambda_{(j)}^{\frac{1}{\alpha}})^{\alpha-1} t^{\gamma}$ as the integrated rate function for channel i. One can show that such a model leads to exactly the same predictions for the distribution of the choice response time and choice probabilities as our dependent horse race model but without assuming dependence among the I processes. However, the underlying multivariate survivor function for this independence model is equal to $S^*_{T_{1(1)},\ldots,T_{1(I)}}(t_1, \ldots, t_I) =$

$\prod_{i=1}^{I} \exp\left(\lambda_{(i)}^{\frac{1}{\alpha}}\left(\sum_{j=1}^{I} \lambda_{(j)}^{\frac{1}{\alpha}}\right)^{\alpha-1} t^{\gamma}\right) = \prod_{i=1}^{I} S^*_{T_{1(i)}}$ and therefore differs from the one in Eq. (7), while giving rise to the same predictions. Thus it seems that there is no process identification because α can also be thought of as a parameter that increases the flexibility of the model without having an interpretation in terms of dependency (as is the case in the independence model).

In order to decide between the dependence and independence model, we need to look for predictions on which the models diverge. Suppose the I single-slot counters correspond to I objects in a choice set. When the effects of removing an object from the choice set are studied (without loss of generality, object I is assumed to be removed), different predictions follow under the two models. For the dependency model, removing object (or channel) I leads to the following overall survivor function:

$$S_T^{(-I)}(t) = \exp\left\{-\left(\sum_{i=1}^{I-1} \lambda_{(i)}^{\frac{1}{\alpha}}\right)^{\alpha} t^{\gamma}\right\}, \tag{13}$$

where the superscript $(-I)$ indicates the removal of object I. However, the independence model (in addition superscripted with an asterisk) leads to the following overall survivor function:

$$S_T^{*(-I)}(t) = \exp\left\{-\left(\sum_{i=1}^{I-1} \lambda_{(i)}^{\frac{1}{\alpha}}\right)\left(\sum_{j=1}^{I} \lambda_{(j)}^{\frac{1}{\alpha}}\right)^{\alpha-1} t^{\gamma}\right\}. \tag{14}$$

It can be seen that the overall survivor functions differ under the two models when an object is removed. (The choice probabilities are the same under both models.) The differential prediction with respect to the response time can be used to distinguish between the dependence and the mimicking independence model.

More generally, for the horse race model based upon a copula approach, there is also a formal identification result. It has been shown by Carrièrre (1995) (see also Escarela & Carrièrre, 2003) that for a horse race model with channel dependencies induced by a copula and prespecified marginal distributions (Weibull distributions), the model is identified.

5.2. Application 2: a parallel-counter model with dependencies

A popular race model is the Poisson parallel-counter model in which it is assumed that the events driving the counters are generated by two independent Poisson processes (Townsend & Ashby, 1983). The simplest model follows by assuming a time-homogeneous Poisson process (with a constant rate), so that the waiting times until a counter reaches its critical count are gamma distributed. The choice and response time are then determined by a race between two independent

gamma-distributed random variables. An extension of this model has been proposed recently by Smith and Van Zandt (2000) who allow the rate of the two Poisson processes to be non-homogeneous or time-varying. However, in the model of Smith and Van Zandt (2000), it is still assumed that the two counters are filled by independently arriving events.

The second application of multivariate counting process is at the same time both an extension and a simplification of the first example. On the one hand, we simplify the model by assuming only two $(I = 2)$ counters. On the other hand, the model is extended by allowing for more than one vacant slot in the two counters. The two counters are denoted as A and B, and they have n_A and n_B vacant slots, respectively. Each counter is associated with a different response (R_A and R_B). As before, the counter that reaches first its criterion determines the response. The random variables T_A and T_B stand for the n_Ath and n_Bth arrival times, respectively. If the counters are independent, the model of Smith and Van Zandt (2000) is obtained.

Suppose now that the two counting processes that fill up the counters are dependent in the way described in this paper. Much of the theory developed by Smith and Van Zandt (2000) can be extended fairly easily to this dependency case. Conditional on the frailty θ, the joint density for the response time and the choice response R_A, denoted as $f_A(t|\theta)$, equals (abusing the probability notation somewhat):

$$
\begin{aligned}
f_A(t|\theta) &= \Pr(T_A = t, T_B > t|\theta) \\
&= f_{T_{n_A}}(t|\theta) S_{T_{n_B}}(t|\theta) \\
&= \frac{\left(\theta\lambda_A^{\frac{1}{\alpha}}t^{\frac{\gamma}{\alpha}}\right)^{n_A-1}\frac{\gamma}{\alpha}\theta\lambda_A^{\frac{1}{\alpha}}t^{\frac{\gamma}{\alpha}-1}\exp\left(-\theta\lambda_A^{\frac{1}{\alpha}}t^{\frac{\gamma}{\alpha}}\right)}{(n_A-1)!} \\
&\quad \times \left\{\sum_{j=0}^{n_B-1}\frac{\left(\theta\lambda_B^{\frac{1}{\alpha}}t^{\frac{\gamma}{\alpha}}\right)^j}{j!}\exp\left(-\theta\lambda_B^{\frac{1}{\alpha}}t^{\frac{\gamma}{\alpha}}\right)\right\}.
\end{aligned}
$$

Integrating over the distribution of θ then gives the joint density of responding R_A and the response time:

$$
\begin{aligned}
f_A(t) &= \frac{\frac{\gamma}{\alpha}\lambda_A^{\frac{n_A}{\alpha}}t^{\frac{\gamma n_A}{\alpha}-1}}{(n_A-1)!}\sum_{j=0}^{n_B-1}\frac{\lambda_B^{\frac{j}{\alpha}}t^{\frac{\gamma j}{\alpha}}}{j!} \\
&\quad \times \int_0^\infty \theta^{n_A+j}\exp\left(-\theta\left[\lambda_A^{\frac{1}{\alpha}}+\lambda_B^{\frac{1}{\alpha}}\right]t^{\frac{\gamma}{\alpha}}\right)g(\theta)\,d\theta \\
&= (-1)^{n_A}\frac{\frac{\gamma}{\alpha}\lambda_A^{\frac{n_A}{\alpha}}t^{\frac{\gamma n_A}{\alpha}-1}}{(n_A-1)!}
\end{aligned}
$$

$$
\times \left\{\sum_{j=0}^{n_B-1}(-1)^j\frac{\lambda_B^{\frac{j}{\alpha}}t^{\frac{\gamma j}{\alpha}}}{j!}\left[\frac{d^{n_A+j}e^{-s^\alpha}}{ds^{n_A+j}}\bigg|_{s=\left(\frac{1}{\lambda_A^\alpha}+\frac{1}{\lambda_B^\alpha}\right)t^\alpha}\right]\right\},
$$
$$(15)$$

where we have applied the special theorem on Laplace transforms used in the proof of Proposition 4 to derive the final expression. The derivation for $f_B(t)$ is analogue. Once $f_A(t)$ and $f_B(t)$ are known, the density of the marginal response time $T = \min(T_A, T_B)$ can be obtained as $f(t) = f_A(t) + f_B(t)$.

The response probabilities $\Pr(A)$ and $\Pr(B)$ follow from integrating $f_A(t)$ and $f_B(t)$, respectively, over t. However, we can make use of a simpler result proven by Smith and Van Zandt (2000) for the case of proportional rate functions. Conditional on the frailty θ, the rate functions of the two Poisson processes driving the counters are indeed proportional. Following Smith and Van Zandt (2000), we may say that the probability of response R_A only depends on the proportionality coefficients $\theta\lambda_A^{\frac{1}{\alpha}}$ and $\theta\lambda_B^{\frac{1}{\alpha}}$ (still conditional on the frailty θ):

$$
\begin{aligned}
\Pr(A|\theta) &= \sum_{j=0}^{n_B-1}\binom{n_A+j-1}{j} \\
&\quad \times \left(\frac{\theta\lambda_A^{\frac{1}{\alpha}}}{\theta\lambda_A^{\frac{1}{\alpha}}+\theta\lambda_B^{\frac{1}{\alpha}}}\right)^{n_A}\left(\frac{\theta\lambda_B^{\frac{1}{\alpha}}}{\theta\lambda_A^{\frac{1}{\alpha}}+\theta\lambda_B^{\frac{1}{\alpha}}}\right)^j \\
&= \sum_{j=0}^{n_B-1}\binom{n_A+j-1}{j} \\
&\quad \times \left(\frac{\lambda_A^{\frac{1}{\alpha}}}{\lambda_A^{\frac{1}{\alpha}}+\lambda_B^{\frac{1}{\alpha}}}\right)^{n_A}\left(\frac{\lambda_B^{\frac{1}{\alpha}}}{\lambda_A^{\frac{1}{\alpha}}+\lambda_B^{\frac{1}{\alpha}}}\right)^j.
\end{aligned}
$$

Because θ cancels from the latter probability formula, the integration over the distribution of θ is not necessary anymore. The mean response times under this model are also easy to calculate because also there θ cancels from the equations. Thus, the formulas obtained under a model with dependencies between the counters are very similar to the formulas presented by Smith and Van Zandt (2000).

To illustrate the effect of the dependency parameter α, we plotted three conditional response time densities in Fig. 2. Decreasing α (i.e., increasing the positive dependency between the two counting processes) shifts the conditional response time density to the left side of the time axis. Moreover, a higher positive dependency counteracts the tendency of the response time densities to become more normal-shaped if the number of vacant slots increases. If $\alpha \to 0$, the probability that the process
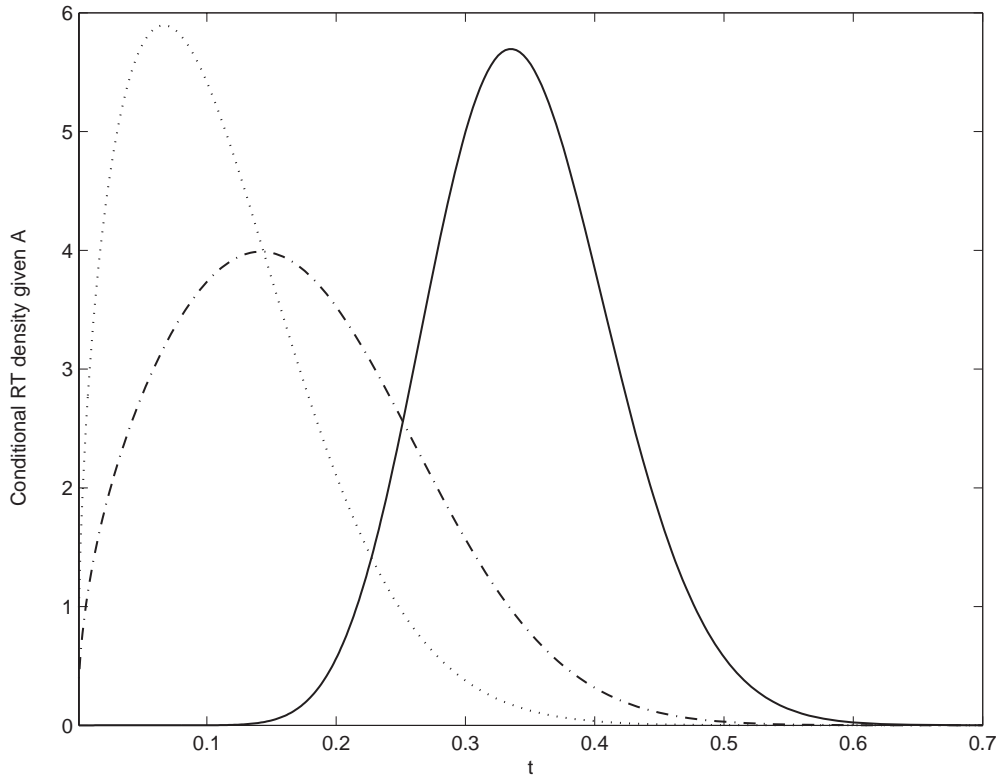
Fig. 2. Three response time densities conditional on responding with option A with common parameters $\lambda_A = 45$, $\lambda_B = 12$, $n_A = 10$ and $n_B = 4$. The dependency parameter varies from $\alpha = 1$ (solid line) to $\alpha = 0.5$ (dashed line) to $\alpha = 0.25$ (dotted line).

with the largest rate function wins the race goes to 1 (unless both rate functions are equal, in which case the dependency parameter has no influence).

The process identification issue discussed in the previous application is also relevant for this example (again, there is no problem on the level of the parameters). As before, we can find an independence model that leads to the same choice probabilities and response times as the dependency model. Let us construct a joint survivor function as follows: $S^*_{T_A,T_B}(t_A, t_B) = (\int_{t_A}^\infty f_A(z)\,dz)(\int_{t_B}^\infty f_B(z)\,dz) = S^*_{T_A}(t_A)S^*_{T_B}(t_B)$. The independence model with underlying joint survivor function $S^*_{T_A,T_B}(t_A, t_B)$ renders exactly the same predictions as the dependency model.

Unfortunately, we cannot rely anymore on a mathematical identification result as in the first application. However, a serious weakness of the proposed independence model is that there is no connection with a realistic underlying counting process. Consequently, the model will have problems in explaining effects of experimental manipulations in a simple way. Assume for example an experiment with two different conditions: one condition with an emphasis on speed and the other with an emphasis on accuracy. Under the dependency model, it is expected that the participants in the accuracy condition raise the number of vacant slots in the counters (n_A and n_B) producing a certain change

in the response time distribution and choice probability. The independence model constructed to mimic the dependency model does not have such an appealing mechanism to explain the observed differences between speed and accuracy conditions because there is no straightforward interpretation of the parameters. The parsimony in explaining results from experimental manipulations may be used to provide evidence to discriminate between the two models.

5.3. Application 3: a coactivation model

Schwarz (1989) and Diederich (1995) proposed a coactivation model based on the superposition of independent homogeneous Poisson processes (also called channel summation) to explain the redundant-targets effect and possible violations of the race model inequality (Miller, 1982). In this last example, an extension of the work of Schwarz (1989) and Diederich (1995) is presented by allowing positive dependence between the summed counting processes (and in addition, we also consider non-homogeneous Poisson processes). Only the double-signal case without stimulus onset asynchrony (SOA equals zero) is discussed. Generalizations to more than two redundant signals and SOAs differing from zero are relatively straightforward.

Assume that a stimulus on redundant trials is presented in two modalities (e.g., visual and auditive) and the task of the participant is to emit a response as fast as possible when the stimulus is detected. Suppose the participant waits for c counts before making a decision. Both modalities drive two correlated channels (A and B) and the events in these channels are summed. Thus, associated with each channel are two counting processes $\{N_A(t), t \geqslant 0\}$ and $\{N_B(t), t \geqslant 0\}$ and the superposition of these two counting processes is defined as $\{N(t) = N_A(t) + N_B(t), t \geqslant 0\}$. Conditional on the frailty θ, the intensity functions in the two separate channels are $\theta \lambda_A^{\frac{1}{\alpha}} t^{\frac{\gamma}{\alpha}}$ and $\theta \lambda_B^{\frac{1}{\alpha}} t^{\frac{\gamma}{\alpha}}$. Applying Proposition 1 shows that, conditional on θ, the channel summation process is again a power function NHPP, but now with intensity function $\theta \left(\lambda_A^{\frac{1}{\alpha}} + \lambda_B^{\frac{1}{\alpha}} \right) t^{\frac{\gamma}{\alpha}}$.

The (unconditional) survivor function of the decision time T_c (the waiting time until the cth event) can be found by using the following equalities $\Pr(T_c > t) = \Pr(N(t) < c) = \sum_{n=1}^{c-1} \Pr(N(t) = n)$. The probability of observing exactly n events is already derived in Proposition 4 for a slightly different intensity function. Considering now $\theta \left(\lambda_A^{\frac{1}{\alpha}} + \lambda_B^{\frac{1}{\alpha}} \right) t^{\frac{\gamma}{\alpha}}$ as the conditional intensity function and integrating over the positive stable density of θ leads to the following survivor function of T_c:

$$\Pr(T_c > t) = \sum_{n=1}^{c-1} (-1)^n \frac{\left[\left(\lambda_A^{\frac{1}{\alpha}} + \lambda_B^{\frac{1}{\alpha}} \right) t^{\frac{\gamma}{\alpha}} \right]^n}{n!}$$
$$\times \left\{ \frac{d^n e^{-s^\alpha}}{ds^n} \bigg|_{s = \left(\lambda_A^{\frac{1}{\alpha}} + \lambda_B^{\frac{1}{\alpha}} \right) t^{\frac{\gamma}{\alpha}}} \right\}.$$

If $\alpha = 1$, the equation simplifies to the one derived in Schwarz (1989). All parameters can be identified uniquely and there are no identifiability problems concerning the dependency character of the underlying process.

6. Conclusion

The derivation of a multivariate counting process with positive dependencies that could be used as a tool to construct realistic mathematical models for cognitive processes was the main goal of this paper. The starting point of the derivation is the relation between the Weibull distribution and the power function NHPP. First, it has been shown that a multivariate Weibull distribution can be constructed using the technique of copulas and that the multivariate Weibull distribution can also be seen as a random-effects model. Next, we

have used this random-effects perspective to create positive dependencies between several power function NHPPs leading to a multivariate counting process. Conditional on the random effect, the individual counting processes are power function NHPPs.

It is shown how the multivariate counting process can be applied to allow for more modeling flexibility in three different counter models: a simple horse race model, a more elaborate parallel-counter model for speeded forced choice tasks and an interactive coactivation model to explain the redundant-targets effect. Because of the tractability of the multivariate counting process, it is relatively easy to find closed-form expressions for the crucial probability densities in the models.

Currently, the important limitation to our approach is that the dependency among the several counting processes is restricted to be positive. Especially for applications where two or more counters are involved in a race, negative dependency would be a useful extension of the independence model. Building a model with negative dependency may also be able to bridge the gap between the two families of sequential sampling models: race models and random walk models. In random walk models, the two counters corresponding to the responses have a perfect negative correlation: Increasing the amount of evidence in one counter with a unit leads to a decrease of a unit in the other counter.

A first possible way to allow for negative dependencies is by not using only survivor functions in the copula but also distribution functions, as suggested by Marshall and Olkin (1988). But, such a construction may lead to very complicated equations. A second method to induce negative associations is by using another copula as the starting point to construct the counting process. However, using other copulas may not lead to equally simple computations as we have obtained for the Gumbel–Hougaard copula, at least not when combined with the Weibull distribution. One could almost say that the Gumbel–Hougaard copula is in a way "conjugate" to first-arrival waiting time distributions that can be derived from several non-homogeneous Poisson processes with proportional intensity functions, with the Weibull distribution as a prominent example. Extending the copula and frailty approach to overcome the positive dependency restriction and using other kinds of distributions will be the topic of future research.

Special care has to be given to the identifiability of the dependency structure in dependent race models that are derived from the multivariate counting processes (the first two examples). It is important to note here that the identifiability issue does not concern the multivariate counting process itself because if the process were fully observed, there would be no problems with the identification of the process. Moreover, the third example in this paper, the interactive coactivation model, is also free of identification problems. For the

identifiability problems in the race models, recent results in the statistical literature show that under some conditions, dependent race models with counters containing only a single vacant slot (i.e., horse-race models) are identified (Abbring & van den Berg, 2003; Carrièrre, 1995; Escarela & Carrièrre, 2003). Some of these results are directly transferable to our model (see the first example where the marginal distributions can be used to identify the model), but unfortunately many of these latter results are not directly relevant for the quite general models considered here (especially not when there are multiple free slots in the counters). However, as has been outlined in the second example, the identification debate can be resolved to some degree by an appropriate setup of the design of the study. The observation that a manipulation of a factor in the research design leads to an expected change in the parameters of the model, adds to the corroboration of the model. An example of this logic is given in the Application section where the parallel-counter model with dependencies is discussed.

Acknowledgments

Appendix A. Simulating a power function NHPP

Simulating the arrival times for n consecutive events in a power function NHPP is straightforward. Let $A(t) = \lambda t^\gamma$ be the integrated rate function. The time of occurrence for the first event is simulated by drawing a realization from a Weibull distribution (using the probability integral transform technique). For the subsequent waiting times, suppose the previous event occurred at time t and we want to simulate the arrival time of the next event. The probability that we have to wait more than s time units until the next arrival is equal to

$$\Pr(N(t+s) - N(t) = 0) = e^{-(A(t+s)-A(t))}$$
$$= e^{-\lambda((t+s)^\gamma - t^\gamma)}.$$

By drawing a realization u for a random variable U distributed uniform on $[0,1]$ and using the inverse survivor function, the arrival time of the next event is

$$t + s = \left(t^\gamma - \tfrac{1}{\lambda}\ln u\right)^{\frac{1}{\gamma}}.$$

Appendix B. The positive stable density

Finding the positive stable density from its Laplace transform is a difficult mathematical problem.

Moreover, it is the characteristic function, and not the Laplace transform, that provides the best starting point for deriving the probability density (see Feller, 1971, p. 548, for a derivation of the positive stable density from the characteristic equation).

There exists another form for the probability density of a positive stable random variable (Hougaard, 2000, Appendix A.3.3), which may lead to simplified calculations.

Proposition 7. *For* $\alpha < 1$, *the probability density from Eq. (9) can be written in the following form:*

$$g(\theta) = \sum_{k=1}^\infty \frac{(-1)^k \theta^{-\alpha k}}{k! \Gamma(-k\alpha)}, \tag{A.1}$$

where the function $\frac{1}{\Gamma(-k\alpha)} = 0$ *for* $k\alpha$ *an integer* (Abramowitz & Stegun, 1974, p. 255).

Proof. Proving this result, we rely on two basic properties of the gamma function: the recursive identity, $\Gamma(1 + x) = x\Gamma(x)$, and the reflection property, $\Gamma(x)\Gamma(-x) = \frac{-\pi}{x \sin(x\pi)}$. Applying first the recursive identity to $\Gamma(k\alpha + 1)$, then the reflection property, and finally inserting the result in Eq. (9) gives the desired result for $\alpha < 1$. □

As said above, it is difficult to find the probability density function of a positive stable random variable from its Laplace transform, but it can be shown that e^{-s^α} is indeed the Laplace transform of Eq. (9) or (A.1).

Proposition 8. *The Laplace transform of the positive stable density* $g(\theta)$ *given by Eq. (A.1) is* e^{-s^α}.

Proof. For the case $\alpha = 1$, the proof is trivial. For the case $\alpha < 1$, we will make use of the following integral identity (Feller, 1971, Eq. (7.9), p. 427), which is repeated here without proof:

$$\int_0^\infty \frac{1 - e^{-s\theta}}{\theta^{y+1}} \, dy = s^y \frac{\Gamma(1-y)}{y}$$
$$= -s^y \Gamma(-y) \quad (y > 0).$$

Next we can compute the following integral transform:

$$\int_0^\infty (1 - \exp(-s\theta)) g(\theta) \, d\theta$$
$$= \int_0^\infty (1 - \exp(-s\theta)) \sum_{k=1}^\infty \frac{(-1)^k \theta^{-\alpha k - 1}}{k! \Gamma(-k\alpha)} \, d\theta$$
$$= \sum_{k=1}^\infty \frac{(-1)^k}{k! \Gamma(-k\alpha)} \int_0^\infty (1 - \exp(-s\theta)) \theta^{-\alpha k - 1} \, d\theta$$
$$= \sum_{k=1}^\infty \frac{(-1)^k}{k! \Gamma(-k\alpha)} (-1) s^{k\alpha} \Gamma(-k\alpha)$$

(using the integral identity)

$$= -\sum_{k=1}^{\infty} \frac{(-s^{\alpha})^k}{k!}$$

$$= -e^{-s^{\alpha}} + 1. \tag{A.2}$$

Now, the left-hand side of Eq. (A.2) can also be written as

$$\int_0^{\infty} (1 - \exp(-s\theta))g(\theta)\, d\theta$$

$$= \int_0^{\infty} g(\theta)\, d\theta - \int_0^{\infty} \exp(-s\theta)g(\theta)\, d\theta.$$

Because the density integrates to one, it cancels against the one on the right-hand side. Multiplying both sides then with -1 proves the proposition. $\square$

Appendix C. Simulating positive stable random variates

An algorithm to simulate from a positive stable density is described by Chambers, Mallows, and Stuck (1976). First draw a value ψ from a uniform distribution on $[0, \pi]$ and next a value w from a standard exponential distribution (with ψ and w independent). Then the quantity

$$\theta = \frac{\sin((1 - \alpha)\psi)^{\frac{1-\alpha}{\alpha}} \sin(\alpha\psi)}{\sin(\psi)^{\frac{1}{\alpha}} w^{\frac{1-\alpha}{\alpha}}}$$

is a draw from a positive stable distribution with parameter α. For $\alpha = 0.5$, the probability density function is displayed in Fig. 1. Note the positive skewness and especially the heavy tails of the distribution: The mass of the graphed part of the density function is about 0.75, thus $\Pr(\Theta > 5) \approx 0.25$. A sample of 100 000 drawings was simulated. The estimated probability of observing a value larger than 100 is still about 5% and the maximum value in the sample is 5.9442×10^8.

References

Abbring, J. H., & van den Berg, G. J. (2003). The identifiability of the mixed proportional hazards competing risks model. *Journal of the Royal Statistical Society, Series B, 75*, 701–710.

Abramowitz, M., & Stegun, I. A. (1974). *Handbook of mathematical functions*. New York: Dover.

Carrièrre, J. F. (1995). Removing cancer when it is correlated with other causes of death. *Biometrical Journal, 3*, 339–350.

Chambers, J. M., Mallows, C. L., & Stuck, B. W. (1976). A method for simulating stable random variables. *Journal of the American Statistical Association, 71*, 340–344.

Colonius, H. (1995). The instance theory of automaticity: Why the Weibull? *Psychological Review, 102*, 744–750.

Cousineau, D. (2004). Merging race models and adaptive networks: A parallel race network. Psychonomic Bulletin and Review, in press.

Cousineau, D., Goodman, V., & Shiffrin, R. M. (2002). Extending statistics of extremes to distributions varying in position and scale and the implications for race models. *Journal of Mathematical Psychology, 46*, 431–454.

Cox, D. R. (1959). The analysis of exponentially distributed life-times with two types of failure. *Journal of the Royal Statistical Society, Series B, 21*, 411–421.

Diederich, A. (1995). Intersensory facilitation of reaction time: Evaluation of counter and diffusion coactivation models. *Journal of Mathematical Psychology, 39*, 197–215.

Dzhafarov, E. N. (1993). Grice-representability of response time distribution families. *Psychometrika, 58*, 281–314.

Escarela, G., & Carrièrre, J. F. (2003). Fitting competing risks with an assumed copula. *Statistical Methods in Medical Research, 12*, 333–349.

Feller, W. (1971). *An introduction to probability theory and its applications*, Vol. II. New York: Wiley.

Fort, A., Delpeuch, C., Pernier, J., & Giard, M. H. (2002). Early auditory–visual interactions in human cortex during non-redundant target identification. *Cognitive Brain Research, 14*, 20–30.

Genest, C., & MacKay, J. (1986). The joy of copulas: Bivariate distributions with uniform marginals. *The American Statistician, 40*, 280–283.

Giard, M. H., & Peronnet, F. (1999). Auditory-visual integration during multimodal object recognition in humans: A behavioral and electrophysiological study. *Journal of Cognitive Neuroscience, 11*, 473–490.

Gumbel, E. J. (1958). *The statistics of extremes*. New York: Columbia University Press.

Gumbel, E. J. (1961). Bivariate logistic distributions. *Journal of the American Statistical Association, 32*, 335–349.

Hougaard, P. (1986a). Survival methods for heterogeneous populations derived from stable distributions. *Biometrika, 73*, 387–396 (correction 75, 395).

Hougaard, P. (1986b). A class of multivariate failure time distributions. *Biometrika, 73*, 671–678 (correction 75, 395).

Hougaard, P. (2000). *Analysis of multivariate survival data*. New York: Springer.

Huber, D. E., & Cousineau, D. (2004). A race model of perceptual forced choice reaction time. In *Proceedings of the 25th annual conference of the cognitive science society*. Mahwah, NJ: Lawrence Erlbaum Associates (in press).

Joe, H. (1997). *Multivariate models and dependence concepts*. London: Chapman & Hall.

Jouini, M. N., & Clemen, R. T. (1996). Copula models for aggregating expert opinions. *Operations Research, 44*, 444–457.

Kochar, S. C., & Proschan, F. (1991). Independence of time and cause of failure in the multiple dependent competing risks model. *Statistica Sinica, 1*, 295–299.

Luce, R. D. (1986). *Response times: Their role in inferring elementary mental organization*. New York: Oxford University Press.

Marley, A. A. J. (1989). A random utility family that includes many of the 'classical' models and has closed form choice probabilities and choice reaction times. *British Journal of Mathematical and Statistical Psychology, 42*, 13–36.

Marley, A. A. J., & Colonius, H. (1992). The "horse race" random utility model for choice probabilities and reaction times, and its competing risks interpretation. *Journal of Mathematical Psychology, 36*, 1–20.

Marshall, A. W., & Olkin, I. (1988). Families of multivariate distributions. *Journal of the American Statistical Association, 83*, 834–841.

Miller, J. (1982). Divided attention: evidence for coactivation with redundant signals. *Cognitive Psychology, 14*, 247–279.

Oakes, D. (1989). Bivariate survival models induced by frailties. *Journal of the American Statistical Association, 84*, 487–493.

Pike, A. R. (1966). Stochastic models of choice behaviour: Response probabilities and latencies of finite Markov chain systems. *British Journal of Mathematical and Statistical Psychology*, *19*, 15–32.

Pike, A. R. (1973). Response latency models for signal detection. *Psychological Review*, *80*, 53–68.

Prudnikov, A. P., Brychkov, Y. A., & Marichev, O. I. (1986). *Integrals and series. Vol. 1: Elementary functions*. New York: Gordon and Breach.

Ritov, Y., Raz, A., & Bergman, H. (2002). Detection of onset of neuronal activity by allowing for heterogeneity in the change points. *Journal of Neuroscience Methods*, *122*, 25–42.

Ross, S. L. (1996). *Stochastic processes*. New York: Wiley.

Sanger, T. D. (2002). Decoding the neural spike trains: Calculating the probability that a spike train and an external signal are related. *Journal of Neurophysiology*, *87*, 1659–1663.

Schwarz, W. (1989). A new model to explain the redundant-signals effect. *Perception & Psychophysics*, *46*, 498–500.

Spiegel, M. R. (1983). *Theory and problems of advanced mathematics for engineers and scientists*. New York: McGraw-Hill.

Smith, P. L. (2001). Stochastic dynamic models (choice, response, and time). In N. J. Smelser, & P. B. Baltes (Eds.), *International Encyclopedia of the Social and Behavioral Sciences* (pp. 15115–15121). Amsterdam: Elsevier.

Smith, P. L., & Van Zandt, T. (2000). Time-dependent Poisson counter models of response latency in simple judgment. *British Journal of Mathematical and Statistical Psychology*, *53*, 293–315.

Townsend, J. T. (1976). Serial and within-stage independent parallel model equivalence on the minimum completion time. *Journal of Mathematical Psychology*, *14*, 219–238.

Townsend, J. T., & Ashby, F. G. (1983). *The stochastic modeling of elementary psychological processes*. Cambridge: Cambridge University Press.

Tsiatis, A. A. (1975). A nonidentifiability aspect of the problem of competing risks. *Proceedings of the National Academy of Sciences*, *72*, 20–22.

Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review*, *108*, 550–592.

Van Zandt, T., Colonius, H., & Proctor, R. W. (2000). A comparison of two response time models applied to perceptual matching. *Psychonomic Bulletin & Review*, *7*, 208–256.