System Identification by Analysis of Failure or Waiting Time Data.

U.Mortensen, U. Suhl, R. Lanwehr

Fachbereich Psychologie und Sportwissenschaften, Institut III,
Westfälische Wilhelms-Universität
48 149 Münster, Fliednerstr. 21
Oktober 1998[1]

Contents

1 Introduction

The aim of the paper is to show how *waiting times* or may be employed either to test hypotheses about the underlying dynamics of psychological or physiological processes, or to arrive at some description of the processes. It will be assumed that a large class of events may be interpreted as resulting from some sort of level-crossing process: a person detects a stimulus, if the corresponding sensory activity reaches a threshold, the person feels depressed, if certain substances resulting from metabolic processes reach a threshold, etc. To this end a hazard-function derived from extreme-value statistics will be presented; this hazard-function allows to define a class of psychometric functions which

[1] Arbeit/art2a/hazard5a2ept2012.tex

are particularly useful if effects of temporal probability summation in detection processes have to be taken into account.

The processes underlying the generation of responses, regardless whether they are responses to health hazards, to events in ordinary life or to stimuli in a laboratory experiment are dynamical processes; unfortunately it is often very difficult to observe these processes directly. Instead, one is often confined to register certain events: the subject responds, a stimulus is detected, a person turns "depressive" or "aggressive", etc. However, the times between such events or until the first occurrence of the event in question may carry relevant information about the processes.

The dynamics of the processes in question are reflected by the time course of certain variables; these are the *trajectories* of the dynamic process. The basic assumption is that the events result if a trajectory hits or exceeds a certain barrier S. The time until the trajectory hits or exceeds S is the *waiting* or *first-passage time*. Generally, the underlying dynamic processes contain stochastic components, so the trajectories may be conceived as sample-paths of some stochastic process. Consequently, the waiting times τ are random variables. The distribution function of waiting times is given by

$$F(t) = 1 - \exp\left(-\int_0^t \phi(\tau)\, d\tau\right) \tag{1}$$

where ϕ is the *hazard function* or *conditional failure rate*, which is related to the conditional probability of the event being observed within the interval $[t + dt)$ under the condition that it has not been observed up to time t (see the next section). Further,

$$G(t) = \exp\left(-\int_0^t \phi(\tau)\, d\tau\right) = 1 - F(t) \tag{2}$$

is known as *survivor function*.

If, on the one hand, the waiting times depend upon some underlying dynamic and if, on the other hand, the distribution function F, or equivalently, the survivor function G depend upon the hazard-function ϕ it is clear that the hazard function reflects in some way the dynamics of the underlying process. To keep things simple we will assume that the occurence of an event of interest depends upon a single variable of the underlying dynamic, represented by some function g. We may then write

$$\phi(t) = \phi(g(t); \alpha), \quad t > 0, \quad \alpha = (\alpha_1, \cdots, \alpha_r)' \tag{3}$$

where α is some vector of parameters.

There are two possibilities:

1. There exists a model of the system and the aim is to identify the model, i.e. to estimate its parameters, or

2. One wants to derive some qualitative features of g from the data (time measurements), perhaps in order to set up a model of the system.

The first case is typically encountered in studies where response times and/or detection responses have to be interpreted, see sections xx and yy. The second case arises typically e.g. in investigations of the effects of some treatment or training, of relapsing-remitting diseases etc. The processes one has to deal with in these situations may be a different kind than those considered for instance in a psychophysical context. The system under investigation may be autonomous and may show, depending upon prevaling conditions,

different types of behaviour: for instance, it may fluctuate around a certain fixpoint or cycle through a series of states. Treatment may simply imply a "disturbance", causing the system to deviate for some time from the closer neighbourhood of a fixpoint and then return to it, in which case the treatment has not really changed anything; the parameters of the system have not been changed. Alternatively, the system may be shifted into the range of attraction of another (stable) fixpoint, or into that of a stable limit cycle. On the other hand, the treatment could imply that the system is shifted away from a limit-cycle, characterising some relapsing-remitting disease, into the range of a stable fixpoint representing a healthier state. All these cases are possible without a change of the parameters of the system, although such a change may be implied by the treatment. In any case, it may be of interest to see how g behaves; in section zz it will be discussed how qualitative features of g may be derived from observed waiting times.

The model considered here differs from the model that is employed in an almost standard way to deal with the effect of external variables, namly Cox's proportional hazard model (Cox (1972)), or a generalisation of this model, like the accelerated life model (Cox and Oakes (1984)). In the proportional hazard model the hazard function is given by

$$\phi(t) = \lambda_0(t) \exp(x'\beta)$$

where $\lambda_0(t)$ is some unknown function, x is the vector of extraneous variables and β a vector of regression weights. The way the x relate to the parameters of the underlying dynamical system is not made explicit. The model is extensively discussed in Kalbfleisch and Prentice (1980); a recent review of work related to this model is given by Henderson (1995). The model proposed in the following allows for an explicit way of formulating the relation between x and the parameters of the model representing the dynamics.

2 General assumptions

Let us briefly recall the defintion of the hazard function. Denote with τ the waiting time for the event. We consider the probability that the event occurs within the interval $[t + \Delta t)$ under the condition that it has not been observed up to time t. This conditional probability may be written as

$$P(\tau \in [t, t + \Delta t) | \tau > t) = \frac{P(\tau \in [t, t + \Delta t) \cap \tau > t)}{P(\tau > t)} = \frac{P(\tau \in [t, t + \Delta t))}{P(\tau > t)}$$

For $\Delta t \to 0$ one has $P(\tau \in [t, t + \Delta t)) \to f(t)dt$, f the density function of the waiting time, and of course $P(\tau > t) = 1 - F(t)$, if F denotes the distribution function of the waiting time. The hazard function is then defined as

$$\phi(t)dt = \lim_{\Delta t \to 0} P(\tau \in [t, t + \Delta t) | \tau > t) \tag{4}$$

and one has immediately from the foregoing

$$\phi(\tau) = \frac{f(\tau)}{1 - F(\tau)} \geq 0, \quad f(\tau) = \frac{dF(\tau)}{d\tau} \tag{5}$$

which implies (1).

We make the following

Assumptions:

1: The event E occurs within the interval $J_t = [0, t)$ if the trajectory $X(t)$ of some stochastic process hits or exceeds the barrier S within J_t,

2: The trajectories are representable as

$$X(t) = g(t) + \xi(t), \quad g(t) = E(X(t)), \tag{6}$$

i.e. g is the mean value function of the process $X_t = \{X(t), t > 0\}$, and $\xi(t)$ is a trajectory of a stationary Gauss-process ξ_t, i.e. $E(\xi(t)) \equiv 0$, $Var(\xi(t)) = \sigma^2 = 1$ and σ^2 is independent of t;

3: The autocorrelation of ξ_t satisfies the condition

$$R(\tau) = 1 + \frac{1}{2}R''(0)\tau^2 + o(\tau) = 1 - \frac{1}{2}\lambda_2\tau^2 + o(\tau), \tag{7}$$

where $\lambda_2 = -R''(0)$ ($R''(0) < 0$ since R has a maximum at $\tau = 0$).

Comments: According to the first assumption we restrict ourselves to events that may be interpreted as resulting from the trajectory of a stochastic process reaching a threshold. According to the second assumption the trajectories are defined by some deterministic function superimposed by noise. This does not necessarily mean that there is a deterministic mechanism generating $X(t)$, although this may be the most straightforward interpretation. The assumption of a stationary noise is certainly restrictive; however, thinking e.g. of the Ornstein-Uhlenbeck-process (OU-process) the assumption is not implausible. The OU-process results from a random walk, representing a particle tossed around randomly, embedded in some fluid exerting friction upon the particle so that it's movement is dampened.

According to the last assumption, the autocorrelation function is assumed to satisfy the conditions for a Taylor-approximation. Since $R(0) = Var(\xi(t))$ one has $R(0) = 1$, and since $R(\tau)$ has a maximum at $\tau = 0$ it follows that $R'(0) = 0$. The assumption of a Taylor-expansion is not compatible with the OU-assumption, because it is not compatible with the Markov-property characteristic for the OU-process. In fact, this assumption has been adopted because it allows for the approximation for the probability of a level-crossing given below. The assumption does not represent a serious restriction of generality because, in a strict sense, Markov-processes represent approximations themselves (Arnold (1973)).

The meaning of λ_2

It is $R(0) = \sigma^2 = 1$, since $R(0) = E(\xi^2(t)) = 1$ according to the assumption. Further, $R'(0) = 0$, since R has in $\tau = 0$ a maximum. To get the meaning of λ_2, consider the process with trajectories $\xi'(t) = d\xi(t)/dt$. $\xi'(t)$ is the rate of change of the function $\xi(t)$ at time t. For $\xi'(t)$ large one has a large change of $X(t)$ at time t, and for $\xi'(t)$ small the change of $\xi(t)$ at time t small. In the extreme case $\xi(t) = constant$ one has $\xi'(t) \equiv 0$. The $\{\xi'(t)\}$ define a family of random functions, since the $\xi(t)$ are random functions of time. So, $\{\xi'(t)\}$ is again a stochastic process. It may be shown (Papoulis (1965), p. 450), that the variance of the $\xi'(t)$, i.e. $E(\xi'^2(t))$, is given by $\lambda_2 = -R''(0)$. On the other hand, the Fouriertransform of $R(\tau)$,

$$S(\omega) = \int_{-\infty}^{\infty} R(\tau)e^{-i\omega\tau}\, d\omega$$

is known as *spectral density function* (*power Density*). $S(\omega)$ characterises the amplitude and phase with which the frequency ω is contained in the trajectories $X(t)$. $R(\tau)$ may

be represented as the inverse transformation of $S(\omega)$:

$$R(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(\omega) e^{i\omega\tau}\, d\omega.$$

Differentiation under the integral yields the derivatives $R'(\tau)$, $R''(\tau)$; for $\tau = 0$ one gets

$$R'(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} i\omega S(\omega) e^{i\omega\tau}\, d\omega, \quad R''(\tau) = -\frac{1}{2\pi} \int_{-\infty}^{\infty} \omega^2 S(\omega) e^{i\omega\tau}\, d\omega,$$

that is

$$R''(0) = -\frac{1}{2\pi} \int_{-\infty}^{\infty} \omega^2 S(\omega)\, d\tau.$$

But the integral defines the second moment of the spectral density $S(\omega)$, so that $\lambda_2 = -R''(0)$ equals the second spectral moment of the process X_t. Since $\lambda_2 = Var(\xi'(t))$, λ_2 is a measure of the speed of fluctuation of the process X_t.

3 Approximations

3.1 Hüsler's approximation

The key idea of deriving the distribution function F, or, equivalently, he hazard function ϕ, is straightforward:

- The interval $J_t = (0, t]$ may be subdivided into n subintervals $(0, t]$, $(t, t_2]$, $(t_2, t_3]$, $\cdots, (t_{n-1}, t_n], t_n = t$, where $t_i - t_{i-1} = \Delta t$ is meant to hold for all i.
- Let X_i^+ be the maximum of $X(t)$ in the i-th interval $(t_{i-1}, t_i]$. Then $X^+ \approx \max(X_1^+, \cdots, X_n^+)$, and for $n \to \infty$, $\Delta t \to 0$ one has

  $$X^+ = \max(X_1^+, \cdots, X_n^+).$$

  For $n < \infty$ one has further

  $$P(X^+ \leq S) = P\left(\bigcap_{i=1}^{n} \{X_i^+ \leq S\}\right) \tag{8}$$

Now (8) is difficult to handle. One reason for this is the discretized time axis. We may thus consider the case $\Delta t \to 0$ with $n \to \infty$ in some appropriate way. However, in this case the dependencies among the X_i^+ can no longer be neglected. However, for sufficiently large values of Δt, where the term "sufficiently large" has to be defined with respect to the speed of fluctuation of $\xi(t)$, the X_i^+ may assumed to be *approximately* stochastically independent (Leadbetter et al (1983)), so that (8) may be replaced by

$$P(X^+ \leq S) \approx \prod_{i=1}^{n} P(X_i \leq S) \tag{9}$$

Because $X_i^+ = g_i^+ + \xi_i^+$ it follows in particular

$$P(X^+ \leq S) \approx \prod_{i=1}^{n} P(\xi_i^+ \leq S - g_i^+) \tag{10}$$

Suppose ξ_t is a Gaussian process. Then the value of ξ at time t is normally distributed, and ξ_i^+ is an extreme value of a normally distributed random variable. As is well known, such values are distributed according to the "double exponential" $\exp(-\exp(-x))$ (Leadbetter et al. (1983)). We may thus surmise that the probabilities $P(\xi_i^+ \leq S - g_i^+)$ are related to the double exponential.

To find the precise limit of (9) for $n \to \infty, \Delta t \to 0$ the dependencies among neighbouring x_i^+ have to be taken into account in an appropriate way.

Theorem 3.1 *(Hüsler's approximation)*[2] *Let ξ_t be a stationary Gaussian process with $E(\xi(t)) = 0$, $E(\xi^2(t)) = 1$ and an autocorrelation function $R(\tau)$ which may be developed into a (Taylor-) series of the form $R(\tau) = R(0)+\tau R'(0)+\tau^2 R''(0)/2+\cdots = 1-\phi_0^2\tau^2/2+o(\tau^2)$. Then the hazard function is given by*

$$\phi(t) \approx \frac{\sqrt{\lambda_2}}{2\pi} \exp\left(-\frac{1}{2}\left(S - g(\tau)\right)^2\right). \tag{11}$$

The proof is a bit lengthy and its details will not be needed in the following; therefore the proof will be given in the appendix.

The survivor function is given by $G(t) = \exp(-\int_o^t \phi(\tau)d\tau)$, so that from (11)

$$G(t) \approx \exp\left(-\frac{\sqrt{\lambda_2}}{2\pi}\int_0^t \exp\left(-\frac{1}{2}\left(S - g(\tau)\right)^2\right)d\tau\right). \tag{12}$$

This is a nice and relatively simple result. The relation of the approximation to the double exponential distribution, induced by the assumption of Gaussian "noise", becomes quite obvious in (12). It has to be kept in mind, however, that the approximation is appropriate only if the value of S is sufficiently large, for instance $S > 2.5$; this follows from the fact that $F(t) = 1 - G(t)$ has been derived within the framework of extreme value statistics. As our extensive simulation studies have shown the approximation (11) is excellent for most practical purposes. Still, there exists another approximation that is not based on extreme value considerations and is therefore of greater generality; it will be briefly presented in the following section.

3.2 Ditlivsen's Approximation

The result of Ditlivsen (1971) will be very briefly outlined, mainly because this result appears to be practically unknown to a wider public, but also because of its usefulness when it comes to evaluate models of dynamic processes. Ditlivsen derived his approximation for the hazard function within the context of reliability questions in engineering and he provided a number of examples from this area. A detailed account is beyond the scope of this paper; it is sufficient to present the main characteristics of his attack.

Let us introduce the abbreviation

$$a(t) := S - g(t) \tag{13}$$

The probability of no level crossing is then given by

$$P(X(\tau) \leq S, \forall \tau \in [0,t]) = P(\xi(\tau) \leq a(\tau), \quad \forall \tau \in [0,t]) \tag{14}$$

[2]We are indebted to Prof. J. Hüsler from the Department of Mathematical Statistics of the University of Berne for his friendly cooperation in this matter; the proof of the following theorem is due to him, and therefore we will speak of *Hüsler's approximation* in the following.

Let us define

$$\Gamma(x,t) := P(\xi(\tau) \le a(\tau), \quad \forall \tau \in [t, t+x]) \tag{15}$$

Now it is well known that for small value of x, i.e. for small intervals $[t, t+x]$, the probability of a level crossing may be approximated by the density function f of $u := \xi(\tau)$ and $v := \xi(t+\tau)$:

$$\begin{aligned} \Gamma(x,t) &\approx P(\xi(t) \le a(t) \cap \xi(t+x) \le a(t+x) + o(x)) \\ &= \int_{\infty}^{a(t)} \int_{-\infty}^{a(t+x)} f(u,v)\, du dv + o(x) \end{aligned} \tag{16}$$

with $o(x)/x \to 0$ as $x \to 0$. The expression (16) may now be related to (15) for the case $t = 0$, $x = T$ in order to arrive at an expression for the probability of a level crossing $1 - P(X \le S)$. The details of the derivations habve to be omitted here:

Theorem 3.2 *Let ξ_t be a wide-sense stationary Gauss process with $E(\xi(t)) \equiv 0$ and $E(\xi^2(t)) \equiv 1$ for all $t \in [0,T]$. Suppose that the autocorrelation function of ξ_t satisfies*

$$R(t) = 1 - \lambda_2 t^2/2 + o(t^2)$$

as $t \to 0$ and that (16) holds. Let $a'(t) = d(S - g(t))/dt = -dg(t)/dt = -g'(t)$, $f(u) = \exp(-u^2/2)/\sqrt{2\pi}$ and $\Phi(u) = \int_{-\infty}^{u} f(t)dt$ and $\gamma = \sqrt{\lambda_2} = \sqrt{-R''(0)}$. Then

$$P(X \le S) \approx \Phi(S) \exp\left(- \int_0^T \left(\gamma f\left(\frac{g'(t)}{\gamma} \right) + \Phi\left(\frac{g'(t)}{\gamma} \right) g'(t) \right) \frac{f(a(t))}{\Phi(a(t))}\, dt \right) \tag{17}$$

The approximation (17) will be referred to as *Ditlivsen's approximation*.

3.3 Some properties of the approximations

To get some intuition about the differences and correspondances between Hüsler's and Ditlivsen's approximation one may consider a special case which shows that differences occur basically only for small values of S, say $S < 3$.

Let P_0 be the probability of a level crossing for the special case $g(t) \equiv g_0$, g_0 a constant, for all $t \in [0,T]$. Then

$$P_0 = 1 - e^{-\Lambda}$$

with

$$\Lambda = \begin{cases} T\sqrt{\lambda_2} \exp(-(S-g_0)^2/2)/(2\pi), & \text{Hüsler} \\ T\sqrt{\lambda_2} \exp(-(S-g_0)^2/2)/(2\pi\Phi(S)) + \log \Phi(S), & \text{Ditlivsen} \end{cases} \tag{18}$$

Hüsler's and Ditlivsen's approximations become equivalent for $\Phi(S) \approx 1$, i.e for $S \to \infty$. For $\lambda_2 \to 0$ one finds for Hüsler's approximation $P_0 \to 0$, while for Ditlivsen's one has $P_0 \to 1 - \Phi(S)$.

If $\Lambda > 0$ it follows from both approximations that

- for $g(t) \equiv g_0$ a constant exponential waiting times are predicted; this is an implication of the assumption that ξ_t is a stationary process;
- for given value of P_0 the values of S and λ_2 cannot be chosen independently.

Although the approximations appear, at a first glance, to be quite different, they lead to identical values for $F(t;S)$ if S is chosen sufficiently large, i.e. for $S > 2.5$, say. For smaller values of S Ditlivsen's approximation still leads to very exact values of $F(t;S)$, while Hüsler's approximation becomes imprecise.

4 Applications

We will briefly indicate how the approximations for the hazard function ϕ can be applied with respect to psychophysical questions or to investigations involving the interpretation of response times. It will then be discussed how they can be employed in studies aiming at a characterisation of a dynamical system underlying observed waiting times.

4.1 Detection experiments

In order to model a detection process one may adopt one of two different postulates: either one assumes that the stimulus activates a certain single channel and, correspondingly, the stimulus is detected by this single channel, or the stimulus activates many channels in parallel and the stimulus is detected if at least one of these channels generates a detection response. We will restrict ourselves to the first assumption, i.e. to the hypothesis of detection by a single channel. Let sss be the presented stimulus and let us further suppose that the channel can, for the stimulus intensities employed, be considered as a linear system. Let g be the response of the detecting sensory system, characterisable by an impulse response $h(t)$. g is given by

$$g(t) = \int_0^\infty h(t-\tau)u(\tau)\,d\tau \tag{19}$$

Let $U(\omega)$ be the Fouriertransform of u und let $H(\omega)$ be the Fouriertransform of h, with $\omega = 2\pi f$, f a (temporal) frequency. In particular the stationary response of the system is given by

$$g_{st}(t) = \int_{-\infty}^\infty H(\omega)U(\omega)e^{i\omega t}\,d\omega \tag{20}$$

While (19) is of interest in particular for inputs u with a rather arbitary temporal structure, (20) is useful if u is sinusoidal, i.e. if $u(t) = m\sin(\omega_0 t)$, where m denotes the amplitude of the input. In this case g_{st} is given by

$$g_{st}(t) = m\,|H(\omega_0)|\sin(\omega_0 t + \varphi(\omega_0)) \tag{21}$$

Now in detection experiments the stimulus is typically presented within a fixed time interval $J_T = [0, T]$. The idea is that the stimulus is detected if

$$X(t) = g(t) + \xi(t) \geq S, \quad \text{for some } t \in J_T \tag{22}$$

This is equivalent to saying that the stimulus is detected if

$$X_+ = \max_{t\in J_T} X(t) = \max_{t\in J_T}(g(t) + \xi(t)) \geq S \tag{23}$$

i.e. if the maximum of $X(t)$ exceeds the threshold S within J_T. Since the trajecory $xi(t)$ is random the maximum X_+ is a random variable; note that in general

$$\max_{t\in J_T} X(t) \neq g(t) + \max_{t\in J_T}\xi(t) \tag{24}$$

Let τ be the first-passage time for the activity of the filter hitting the barrier S, and let $F(t) = P(\tau \leq t)$ be the distribution function of τ. Now $X_+ > S$ if $\tau \leq T$, so that one has

$$P(X_+ > S) = P(\tau \leq T) = F(T) \tag{25}$$

Consequently, the distribution function for X_+ is given by the distribution function for the first-passage time,

$$P(X_+ \leq S) = 1 - F(T) \tag{26}$$

Also, the psychometric function $P_\psi(m)$, defining the probability of detection for the amplitude or contrast m of the input u is given in terms of F:

$$P_\psi(m) = P(X_+ > S|m) = 1 - F(T;m) = G(T|m) \tag{27}$$

where $F(T;m)$ instead of $F(T)$ was written in order to indicate that m is considered to be the relevant parameter now. Actually,

$$P_\psi(m) = G(T;m)$$

is now considered a function of m.

At this stage, there are again two possibilities: (a) there exists sufficient knowledge about the detecting system to assume a particular form for the impulse response h, and (b) this knowledge does not exist, so that the shape of h has to be derived from the data. The situation (b) was considered by Roufs and Blommaert (1980). They suggested there *perturbation method*, which is based upon the assumption

$$\max_{t\in J_T} X(t) = \max_{t\in J_T} g(t) + \max_{t\in J_T} \xi(t) \tag{28}$$

This assumption is well known under the name *peak detection*. In view of the fact that in general (24) is true the peak detection assumption (28) appears to represent rather a special case; indeed, it takes only little thought to see that (28) can be true in the strict sense only if the $\xi(t)$ do not vary within a trial and change only between trials. This appears to be a very restrictive condition on the noise process ξ_t. On the other hand, Roufs and Blommaert (1980) determined the impulse and the step response of the detecting system (either of sustained or of transient type) and were able to predict the empirical step response from the empirical impulse response, employing (28). This means that the postulate of peak detection seems to hold as an approximation, at least under the condition of the experiments of Roufs and Blommaert (1980). We may suspect that the approximation holds either if the value of λ_2 is very small, meaning there exist only very slow random fluctuations, or S has a large value. i.e. the high-threshold assumption is adequate, and λ_2 is correspondingly large, so that the threshold is reached only in close neighbourhood of the maximum of the impulse response. Indeed, Roufs and Blommaert (1980) observed only a very small number of false alarms, justifying a large value of S and a corresponding large value of λ_2. This was verified in a simulation study by Mortensen and Suhl (1990) where it was shown that λ_2 assumes values compatible with the common assumption that the noise in a sensory system is practically white.

If, on the other hand, a model for h is considered, i.e. if a particular function is assumed for h, then the parameters S, λ_2 and those of h may be estimated employing standard procedures like maximum-likelihood; this will not be pursued here. Instead, another look at detection data appears to be worth while. Consider an experiment where the stimulus amplitude is varied sinusoidally with time. If the detecting system is linear, the response is given by (21), i.e.

$$g_{st}(t;m) = m\,|H(\omega_0)|\sin(\omega_0 t + \varphi(\omega_0))$$

Usually one assumes that the stimulus will be detected with probability $p_0 = .5$, say, if

$$\max g_{st}(t;m) = c,$$

where c is some constant. Now the response will be maximum when

$$\sin(\omega_0 t + \varphi(\omega_0)) = 1$$

so the stimulus is detected if

$$m(\omega_0)|H(\omega_0)| = c$$

which again implies that the "sensitivities" $1/m(\omega_0)$ are proportional to $|H(\omega_0)|$, i.e.

$$\frac{1}{m(\omega_0)} \propto |H(\omega_0)| \tag{29}$$

So, if (29) holds one may empirically determine $|H\omega_0)|$ by determining the value of $m(\omega_0)$ for sufficiently many values of ω_0. Unfortunately, the length T of the observation interval does not enter here. Suppose that T is held fixed for various values of ω_0; since the number of peaks, i.e. the number of times at which $g(t) = m|H(\omega_0)|$ varies with ω_0, the effects of $H(\omega_0)$ (low-pass or band-pass) will be confounded with the increase of detection probability with the number of peak values. To counteract this confoundation one may vary T with ω_0 such that T is determined by the number of complete cycles of the sinusoid. This was done by Roufs (1974). He found that for given value of ω_0, $m(\omega_0) = m(\omega_0; k)$ depends upon the number k of cycles according to

$$\log m(\omega_0; k) = \alpha \log k + \beta, \quad \alpha < 0 \tag{30}$$

The value of α turned out to be independent of ω_0, but β does depend upon ω_0. Employing this relationship together with the fact that the Crozier-coeficient is constant (= .25) for different the values of ω_0, S, λ_2 and $|H(\omega_0)|$ may be estimated from the data using again Hüsler's or Ditlivsen's approximation. Both approximations yield the same results, so it is sufficient to restrict oneself to the less computationally expensive Hüsler-approximation. Since the stimuli were gated the on- and off-transients may be neglected, so that

$$\int_0^T \phi(\tau)\, d\tau = k \int_0^{2\pi} \phi(\tau)\, d\tau$$

where $g(t) = m(\omega_0)|H(\omega_0)| \sin(\omega_0 t)$ is inserted; note that the phase $\varphi(\omega_0)$ need no longer be considered. The estimates for S and λ_2 correspond to those derived from the impulse response data; details may be found in Mortensen and Suhl (1990).

4.2 Response times

Suppose a stimulus is presented and the task of the subject is to respond as quickly as possible. The response times may be of interest because they may contain information about sensory processes as well as on cognitive processes. It should be noted that the approximations of Hüsler (11) on the one hand and of Ditlivsen (17) on the other provide models for response times. Although a vast variety of experimental paradigms for the collection of response times and models for their analysis have been suggested (Luce (1986)), models based on level crossings as considered here do not seem to have been proposed so far. This may stem from the fact that explicit expressions relating a system's response - g - to the probability of a level crossing are not to be found in the literature on stochastisc processes, except for cases which are not of interest for the modelling of response times (e.g. $g = \text{const}$, see Cramér and Leadbetter (1967)). Equally likely, however, certain traditions in interpreting response times may play a role. Donders (1868) suggested that the completion of a task may be decomposed into components and that each component requires a randomly distributed time. According to Donders' model,

the total (observed) response time is the sum of the times required for the components. Processing of the stimulus and arriving at a decision about when to respond is thus conceived as a serial process. Alternatively, one may consider the organism as a bank of parallel systems ("channels"); the subject responds as soon as one of the channels responds. The response time is then the minimum of the times required by the channels to process the stimulus. Of course, all sorts of mixtures of serial and parallel processing models are conceivable. The key assumption for all these models is that the total response time somehow depends upon the times needed for the components to process the stimulus.

In its generality this assumption can hardly be disputed. The question is, *how* the components contribute to the total response time. That the times required by the components to process the stimulus will dependent upon each other appears to be evident if one dares to have a look at real systems. Whether the structural composition of the system on the basis of components which are not specified any further may be deduced from decompositions of total response times may reasonably be doubted.

The brain is a network of interlinked neurons and it is not quite clear how the above mentioned components should be defined with respect to the network. It is natural to think of the activation of parts of the network in terms of diffusion processes (e.g. Murray (1989)); a variety of models relating responses times and processing of stimuli in terms of diffusion processes is discussed in Luce (1986); c.f. also Diederichs (1992). The model proposed here may be considered as an approximation to a diffusion model. Indeed, let L be the system of neurons that processes the stimulus. L is assumed to be deterministic and the noise to be additive and stationary. This assumption is clearly a simplification, but in the light of its success in catering for detection data it may be a useful one. Note that L may encompass parallel and serial components. In general, L will be nonlinear, although under certain circumstances a linear approximation of L may do (c.f. section 4.1). The great advantage of (11) and (17) is that only L, having the output g, has to be modelled, and nothing has to be assumed about the times required by individual components of the system. Even the motor response component may be modelled to be contained in L.

In the following we concentrate on what is considered *simple reaction times*. The reaction time (RT) is assumed to be decomposable according to $RT = R + D$, where D is the *decision time* and R the *residual time*. The decision time D is the time psychologists seem to be most interested in; this is the time required by cognitive processes to arrive at a decision whether to respond or not. The residual time R reflects the sensory processing and the motor response. For instance, in the experiments of the type performed by Roufs and Blommaert (1980) one is primarily interested in the identification of mechanisms responsible for sensory processing. In the level-crossing approach considered here – the stimulus is detected when the activity reaches a threshold value S – the decision process is incorporated in the model: it is assumed that the subject decides that a stimulus was presented whenever the trajectory $\xi(t)$ representing the activity reaches S. The value of S may have been "set" implicitly by cognitive processes which again is governed by the experimental task and setting. So what is missing in order to make (11) a complete model for simple reaction times is a component describing the motor response. It may be fruitful to determine the (random) times required for the motor response once they are triggered by $\xi(t)$ reaching the value S. In any case, the response time cannot be conceived as being given by the convolution of R and D.

In Fig. 1 the distribution function for the reaction times as generated by (11) are shown for the sustained channel as determined by Roufs and Blommaert (1980). Here $\Lambda_2 = \sqrt{\lambda_2}/2\pi$. Two distributions are shown, one for a contrast of $m = .015$ and one for $m = .1$. The distributions differ slightly; the amount by they differ depends upon the

Figure 1: Distribution functions and impulse response of a sustained channel

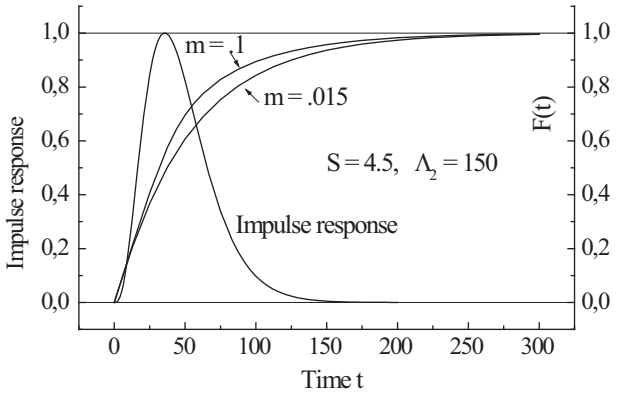
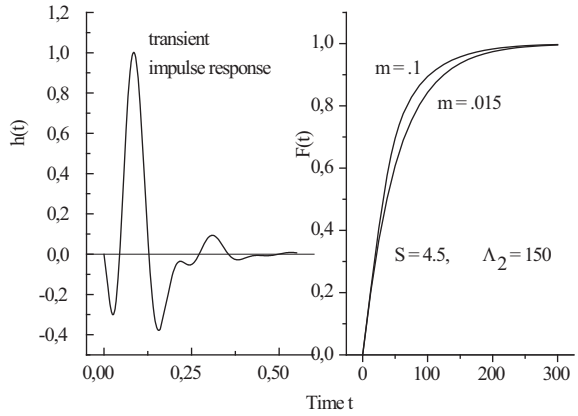


Figure 2: Transient impulse response and distribution functions

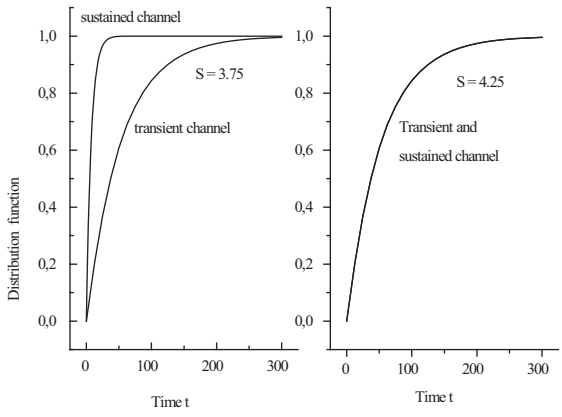


value of S. Fig. 2 shows the corresponding distributions for the transient channel. The distribution functions differ for the two types of channels differ with respect to the time scale;

Discuss Roufs and v.d. Brincke and their model shortly here!

It may be of interest to have a brief look at the response times for sinusoidally flickering stimuli. If the amplitudes are within threshold range g will be sinusoidal as well and the distribution functions have the form depicted in Fig. **??**; ϕ is the hazardfunction, not g. If the amplitudes are above threshold, g will no longer be sinusoidal because of the nonlinearity of the sensory system. The distribution function will retain there staircase like character, the precise form depending upom the nonlinear distortion of g. The discussion in the following section will provide a feeling for oscillating processes from nonlinear system.

Figure 3: Distribution functions for transient and sustained impulse-responses, different S-values



5 Estimating qualitative aspects of the mean value function $g(t)$.

5.1 General considerations

Let us suppose that no specific model for g is given and that one wants to derive some qualitative impression of g from the data. We assume that the data are waiting times without any censoring, the systems starts always in the close neighbourhood of a certain state characterised by certain starting values of the trajectories of the system. Data of this sort may arise from observations of a single individual.

Suppose that some estimate of the survivor function

$$G(t) = \exp\left(-\int_0^t \phi(u)\,du\right)$$

is available. For sufficiently small Δt one has from Taylor's expansion

$$G(t+\Delta t) \approx G(t) + \Delta t G'(t) \tag{31}$$

with

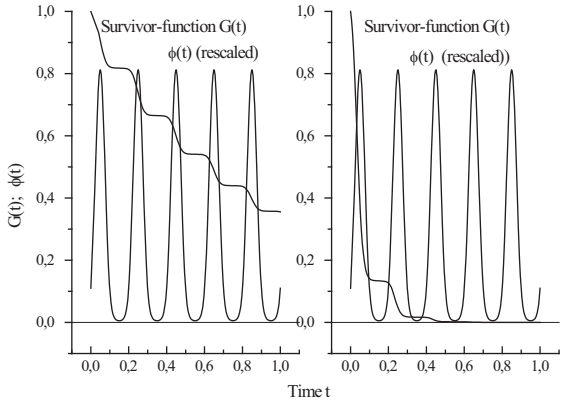
$$G'(t) = \frac{dG(t)}{dt} = -\phi(t)G(t).$$

Solving for ϕ and substituting e.g. Hüsler's approximation one finds after some straightforward algebraic transformations

$$g(t) - S \approx -\sqrt{-2\log\frac{G(t)-G(t+\Delta t)}{A\Delta t G(t)}}, \quad A = \frac{1}{2\pi}\sqrt{\lambda_2} \tag{32}$$

The r.h.s. of (32) yields an estimate of g if an estimate $\hat{G}$ of G is given, e.g. as a smoothed version of a Kaplan-Meier estimate, and if the values of S and $A = \sqrt{\lambda_2}/2\pi$ are known. Of course, the values of these parameters will in general not be given. Now the survival function G does depend upon S and A, and so will any estimate of G. Suppose that, in order to apply (32), we assume an arbitrary value of A, for instance $A = 1$. For this particular value of A the r.h.s. of (32) yields a first estimate of $g(t) - S$; the fact that

Figure 4: Survivor-functions for flicker stimuli for different values of λ_2



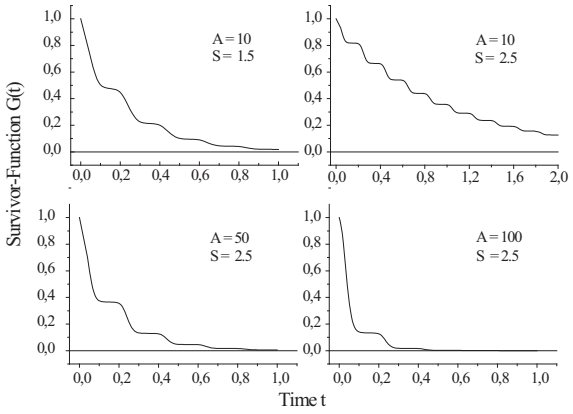
we do not now explicitly the value of S does not really matter as long as we are only interested in the *shape* of g, because g and $g - S$ differ only by the constant S. The question is thus to what an extent the estimate of the shape of g is distorted by the arbitrarily chosen value $A = 1$ (or $\lambda_2 = 1$). As will be demonstrated in the following subsection the choice $A = 1$ seems to be a lucky one: regardless of the true value of λ_2 this choice of the value of A appears to imply very little distortion of the shape of g.

However, what does play a role for the estimation of g is the relation between the *true* values of S and λ_2: if S is small compared to that of λ_2 the waiting times t_i will, in general, be shorter than the time needed for g to reach its first maximum so that only this part of g can be recovered from the t_i and any periodicities in g will remain undetected. In Figure 4 the shape of the survivor function G is plotted together with the hazard function $\phi(t)$ for $g(t) = \sin(\omega t)$. G decreases in a staircase-like manner, where a single step corresponds to a cycle of g as reflected by ϕ. The width of the plateaus corresponds to the width of the valleys of ϕ. If the value of λ_2 is "large" compared to that of S the number of steps is reduced possibly to zero. So, if no periodicity is observed this may be due to the value of λ_2 being large compared to that of S. To the extent the value of S can be manipulated experimentally, for instance by choosing a more restrictive definition of the events in question one may try again with a higher value of S: for each value of $\lambda_2 < \infty$ one can find a value of S such that the periodicity of g, if it exists, can be detected.

The particular combination of values for S and λ_2 underlying the data is vital for recovering the course of g from that of $\hat{G}$. Figure 4 shows two survivor-functions G, corresponding to different values of $A = \sqrt{\lambda_2}/2\pi$, together with a plot of $\exp(-(S - g(t))^2)2)$; for $A = 10$ one has $\lambda_2 = 628.32$ and for $A = 100$ this parameter assumes the value $\lambda_2 = 62831.85$, i.e. the second spectral moments differ by a factor of 100. Fig. 5 illustrates the point implicit in Fig. 4 again. Four survivor-functions corresponding to different values of S and $A = \sqrt{\lambda_2}/2\pi$. A decrease of S for given value of A or a decrease of A for a given value of S implies that G approaches 0 (i) faster and (ii) shows a smaller number of "plateaus" which indicate the oscillatory character of g. An "unpleasant" case is the one characterised by $A = 10$, $S = 2.5$: here the steps are so small that they are likely to be buried in the steps that one has in estimates of G as given by the Kaplan-Meier estimator anyway; smoothing the estimate will be a vital exercise in such cases.

Now A reflects the speed with which the noise fluctuates, and this is an aspect of the process the experimenter most likely cannot manipulate. The definition of the observed

Figure 5: Survivor functions for different values of A and S



event, however, contains implicitly a definition of the value of S. So if only a small number of plateaus is observed (if any of them are observed at all) one could try to define the events more restrictively, implying a larger value of S, in order to increase the number of "steps" in the observed survivor-function; we will take up this point in the Discussion.

5.2 Activator-Inhibitor-systems

In this section we present some simulations showing that it is, in principle, possible to recover the function g underlying the distribution of waiting times. To this end we consider autonomous systems showing cyclic behaviour. Such systems are necessarily nonlinear; a biologically interesting class of systems is that of activator-inhibitor systems. Perhaps the simplest system of this type is the Lotka-Volterra-system modelling the relation between a prey and a predator population. In standardised form (vergl. Murray (1989), p. 64) as

$$\frac{du}{dt} = u(1-v), \quad \frac{dv}{dt} = \alpha v(u-1) \tag{33}$$

Figure 6 shows an example for $u(t)$ and $v(t)$ for $\alpha = .5$. Again the "noise" may be additive and independent of $u(t)$ and $v(t)$. Suppose $g(t)$ is defined by one of the variables u or v; let $g(t) = u(t)$. The extent as to which the shape of g can be estimated from the data was investigated by simulation studies. To this end random times were generated on the basis of Hüsler' approximation

Fig. 7 shows (simulated) data points as they may be observed; the points are shifted with respect to the true value of S. A shift corresponding to the true value shows that the points do indeed correspond to the mean value function $g(t)$ as defined by the Lotka-Volterra-model very well. Although the value of λ_2 is rather large the (true) value of S is as well. Note, however, that the "valleys" of g are not represented by the estimated points. So if data like these are actually observed one may surmise that g is defined by a system like the Lotka-Volterry and try to estimate the parameters of the model so that a fit like that in Fig. 8 results.

The Lotka-Volterra system has the advantage of being very simple. However, it is also known to be implausible: in the absence of the inhibitor the activator will grow indefinitely. As a result, the system is known to be *structurally instable*: the slightest

Figure 6: The interaction of an activator and an inhibitor: the model of Lotka-Volterra

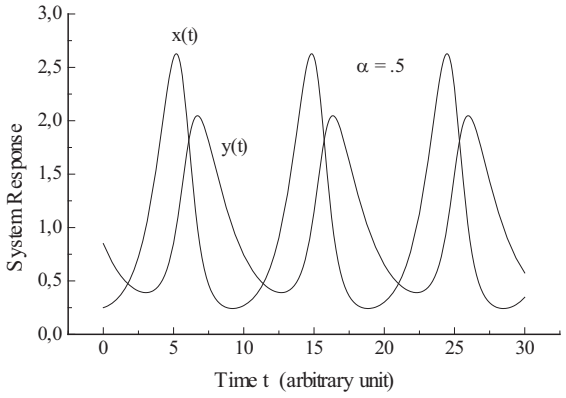
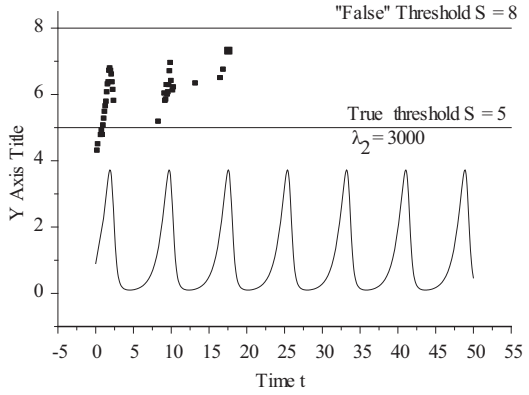


Figure 7: Estimates of $g(t) = u(t)$ from a Lotka-Volterra-system (I)



disturbance of the system causes a change of the trajectory, i.e. there is no return to some stable state or limit cycle. In the following modification of the original Lotka-Volterra-system is in many ways more realistic.

Suppose the (activator-inhibitor-) system is defined by

$$\frac{du}{dt} = f(u,v) = u(1-v) - \frac{auv}{u+d} \tag{34}$$

$$\frac{dv}{dt} = h(u,v) = bv\left(1 - \frac{v}{u}\right) \tag{35}$$

The free parameters are a, b and d; they are dimensionless and reflect the relative effects of the actual parameters of the system (cf. Murray (1989), p. 73); for our purposes, the real parameters are not of interest. This system may have stable and/or instable equilibrium states or stable and/or instable limit cylces. The system is in equilibrium if $du/dt = dv/dt = 0$. Suppose this occurs for $u = u^*$, $v = v^*$, so then $f(u^*, v^*) = 0$ and $h(u^*, v^*) = 0$. It follows from (35) that then

$$u^*(1-u^*) - \frac{au^*v^*}{u^*+d} = 0 \tag{36}$$

Figure 8: Estimates of $g(t) = u(t)$ from a Lotka-Volterra-system (II)

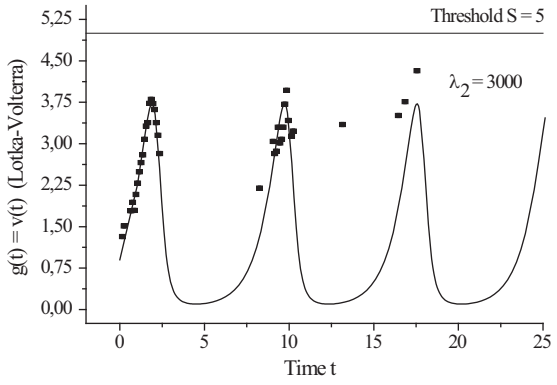
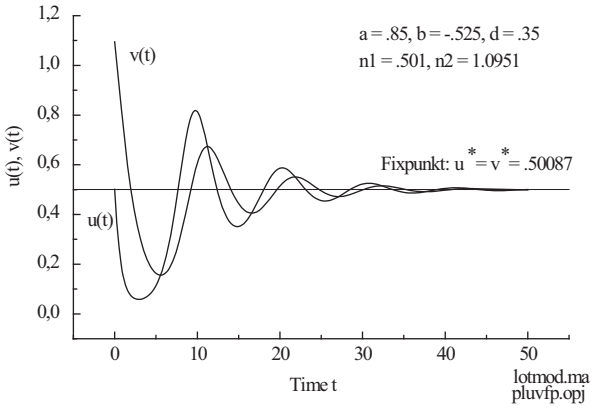


Figure 9: Approach to fixpoint of system



$$bv^* \left(1 - \frac{v^*}{u^*}\right) \quad = \quad 0 \tag{37}$$

From (37) it follows that

$$u^* = v^*, \quad u^{*2} + (a + d - 1)u^* = 0$$

define the solutions. If one concentrates on positive solutions (because only these may make biological sense) one has

$$u^* = \frac{(1 - a - d) + \sqrt{(1 - a - d)^2 + 4d}}{2} \tag{38}$$

The next question to be tackled refers to the stability of the states. To this end, the system will be linearised at the equilibrium points. The details will be omitted here; it is sufficient to state that there are two domains of parameter values: a stable and an unstable one. The system is stable for all $0 < a < 1/2$ and $b > 0, d > 0$. Within the instable range there exists a subrange of parameter values such that the trajectories of the system converge towards a limit cycle, i.e. the system performs some periodic movement. For instance, for $a = .15$, $b = -.125$, $d = .5$ and for the start values

Figure 10: Approach to the fixpoint of a system

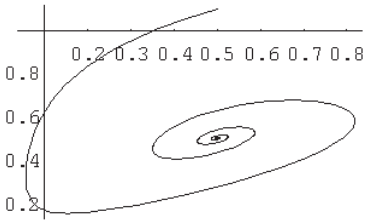
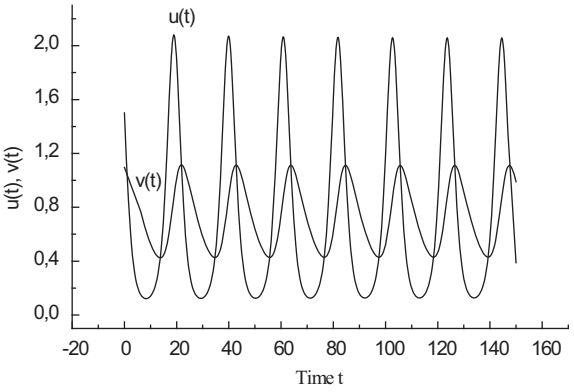


Figure 11: Approach to a stable limit cycle



$u(0) = 1.501$, $v(0) = 1.0951$ the system approaches a stable fixpoint. If the value of a is changed to $a = .375$ the system enters a stable limit cycle, see Figure 12.

We were interested to see to what extent periodic behaviour of $g(t)$ can be recovered from waiting times. To this end the approach to a stable fixpoint, as shown in Figure 10, could have been chosen for a simulation study; indeed, this case is of interest because this could be the situation at hand, for instance after some therapeutic steps have been taken. On the other hand an overall decrease of the values of $g(t)$ as in Figure 10 could yield a confounded image of the power of our method to reconstruct the function g, so that we decided to stick to parameter values for which the system shows limit cycle behaviour. As already mentioned one generally does not know the values of S and λ_2. Figures 13, 14, and 15 show the results for three different cases: in Figure 13 the data were generated assuming $S = 4$ and $\lambda_2 = 1000$; for the plot (a), however, $\lambda_2 = 1$ was assumed. When the plot is superimposed the true function g one sees that the important overall features of g are caught by the data. As Figures 14 and 15 show this holds also if the true value of the second spectral moment is given by $\lambda_2 = 2000$ and $\lambda_2 = 5000$. Of course, the smaller the value of λ_2 for given value of S the better; Figure 16 shows the case where the true value of λ_2 is $\lambda_2 = 500$.

Suppose again that $\lambda_2 = 1000$, but that only $n = 20$ data points are sampled; then the picture given in Figure 17 results. These values do not really provide any information about g at all. If the number of observations is increased to $n = 100$ the situation does improve, as may be seen from Figure 18: clearly, it becomes obvious that g is in some form periodic. It seems that a minimum of 150 measurements is necessary to get a feeling of the shape of g.

Figure 12: Approach to a stable limit cycle

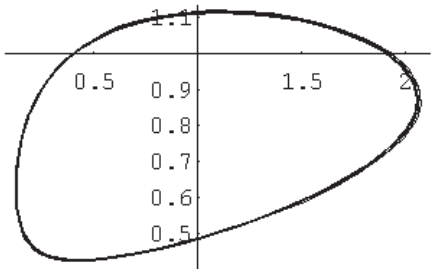
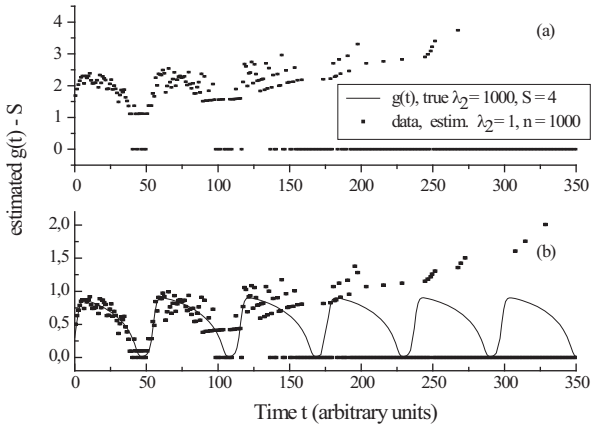


Figure 13: b7esttru



6 Discussion

The aim of the paper was to present a model that relates waiting times to trajectories of a dynamical system. The stochastic aspects of waiting times are catered for assuming additive, stationary Gaussian "noise" ξ_t, characterised by the second spectral moment λ_2. This parameter reflects the speed of fluctuations of the process ξ_t. A low value of λ_2 means slow fluctuations, and thus high autocorrelations, a high value fast fluctuations and quickly decreasing autocorrelations. Here we have a first advantage of our approach: in contrast to time series models like the one proposed by Albert et al. (1994) for relapsing-remitting diseases just two parameters are required to characterise the stochastic aspects of the process at hand, namely λ_2 and the threshold S. If λ_2 assumes a large value, S has to be "large" in a corresponding way in order to allow for either an estimation of the parameters of g, characterising the system, or of some qualitative evaluation of g. The parameter S is again linked to the specification of the events to be observed.

The model is certainly useful when one is interested in the interpretation of response times or detection data. Here one may assume that at time $t = 0$, say, the dynamics are started in an identical way, apart from stochastic aspects. When one is dealing with response times the motor component has to be modelled additionally. In any case, one may consider explicit models of the processes of interest without having to resort to general purpose models of the hazard function, like the Cox-proportional-hazard function

Figure 14: b8esttru

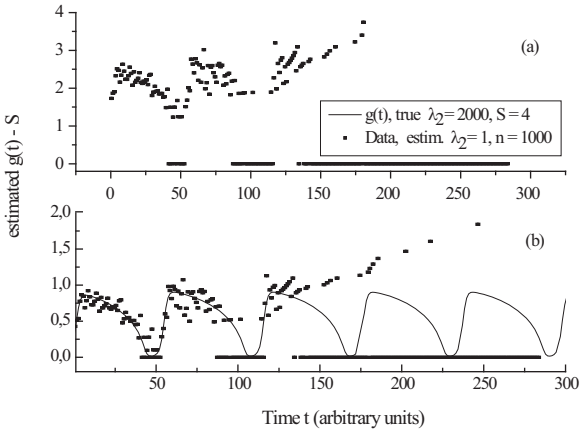
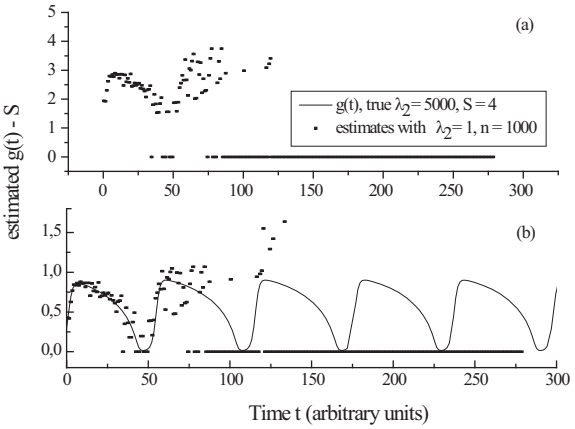


Figure 15: b9esttru



or generalised versions of it, e.g. the accelerated life model.

Our model allows also, at least in principle, to evaluate other types of processes. For instance, one may be interested in the course of relapsing-remitting diseases, like certain forms of depression, or multiple sclerosis, where regular measurements of relevant variables are not possible or simply not available, and where times between certain events relating to the process are given instead. The underlying process may be characterised in various ways:

1. The system is passive, i.e. $g(t) \equiv 0$ if no there is no appropriate stimulation. This case may be assumed e.g. in psychophysical and/or response time studies.

2. The variable of interest may oscillate in a regular fashion, and the system is autonomous in the sense that there is no external driving force oscillating in the observed way. The oscillations may reflect a limit cycle, the approach to a limit cycle, or simply the approach to a fix point; which of these three possibilities has to be assumed is difficult to decide since it seems to be possible only to recover the

Figure 16: b10akomp

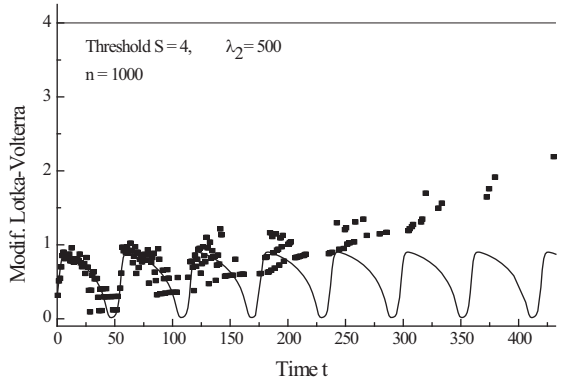
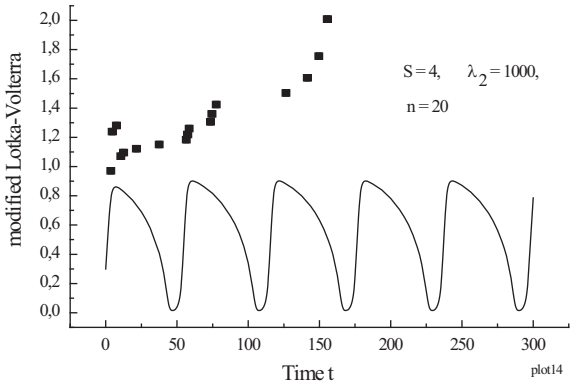


Figure 17: Plot14



first two or three cycles.

3. The system is autonomous, but very damped: if conditions are changed it approaches a corresponding fix point.

Autonomous oscillating systems are necessarily nonlinear. Dealing with such systems offers new ways of looking at data, or at the effect treatments may have. For instance, the system governing the behaviour at hand may have several stable fix points. The actual behaviour of the system consists in fluctuating around such a fix point, due to some stochastic disturbances. For instance, the fixpoint may characterise a depressive mood. Therapy may now act in two ways: (i) it changes the system parameters, like a, b and d in (35), or (ii) leaves the parameters invariant but simply shifts the system into a different potential dwell. In this case nothing is really changed, appropriate random shocks may shift the system back into the depressive state. In case the parameters are changed the system may also be shifted back into a depressive state by random shocks from the outside (or inside) world, but the depressive potholes may be not as deep, so the systems is pushed out of them more easily, behaving "non-depressed" in general.

Interesting as these aspects of the model may be, the question of estimating g is thorny. The estimation of g requires decent estimates of the survivor function $G(t)$ for a

Figure 18: Plot18

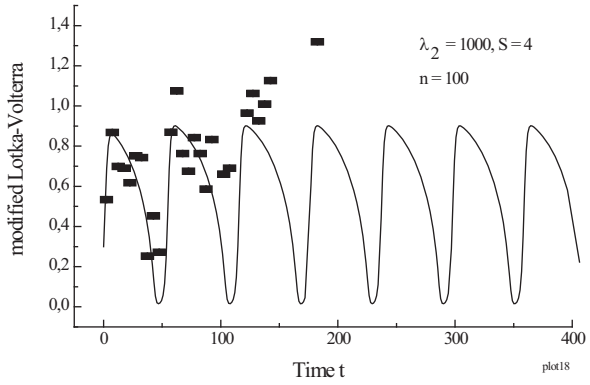
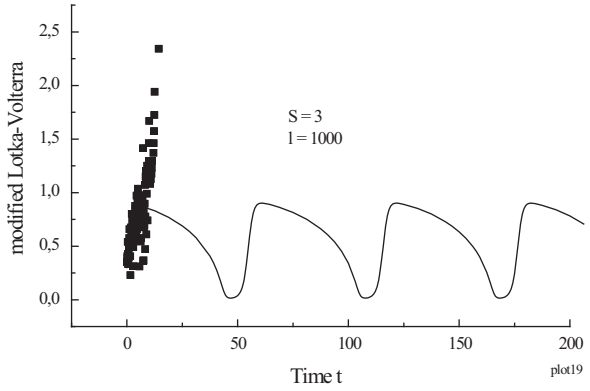


Figure 19: Plot19



start; one will have to discuss smoothed versions of Kaplan-Meier-estimates here. With regard to the discussion of response times the problem has been discussed e.g. by Bloxom (1985a, 1985b). With regard to a study about depression prevention a different line of attack was chosen by Greenhouse and Silliman (1996), considering in particular certain problems arising from censoring.

Apart from this one has to deal with the question concerning the relation between the values of the threshold S and that of the second spectral moment λ_2; if S has a small value compared to that of λ_2 the observed times will not reflect more than the first increasing flank of g, so that not much information about the sytem generating g can be gained from the data. Let us consider a case where the relation between S and λ_2 is not optimal with respect to our purposes; this is for instance the case when the value of S is too small relative to that of λ_2. Such a situation is given in Figure 19. The data do no longer reveal the underlying periodic g: the waiting times do not exceed the time at which the first maximum of g occurs.

7 Proof of theorem 3.1

Let the interval $(-a, a)$ be subdivided into n subintervals of length Δs, and let $Y_i = \max_{(i-1)\Delta s, i\Delta s} Y(x) = Y_i(\Delta s)$. For sufficiently large $S - g_i$ one has

$$P(Y \leq S) \approx \prod_{i=1}^{n} P(Y_i \leq S) \tag{39}$$

(asymptotic independence).

Let $\mathbb{E}(N_u(1))$ be the expected number of Up-Crossings for an arbitrary level u per spatial unit. Then the *Rice-formula* is

$$\mathbb{E}(N_u(1)) = \frac{\sqrt{\lambda_2}}{2\pi} \exp(-\frac{u^2}{2}) := \mu(u) \tag{40}$$

(cf. Leadbetter, Lindgren und Rootzen (1983), p. 153). Let

$$\Phi(u) = \frac{1}{2\pi} \int_{-\infty}^{u} \exp\left(-\frac{z^2}{2}\right) dz$$

Further, according to Leadbetter, Lindgren and Rootzen (1983), p. 166, Lemma 8.2.1:

1. For all $\Delta s > 0$ one has

$$P(Y_i(\Delta s) > u) \leq 1 - \Phi(u) - \mu(u)\Delta s \tag{41}$$

2. For $\Delta s > 0$ and for $\Theta < 1$ there existists a $\Delta s_0 = \Delta s_0(\Theta)$ such that for $0 \leq \Delta s \leq \Delta s_0$ the inequality

$$P(Y_i(\Delta s) > u) \geq 1 - \Phi(u) + \Theta\mu(u)\Delta s \tag{42}$$

holds.

The inequalities may be summarised in the form

$$1 - \Phi(u) + \Theta\mu\Delta s \leq P(Y_i(\Delta s) > u) \leq 1 - \Phi(u) + \mu\Delta s. \tag{43}$$

For $u \to \infty$ and the fact that Θ may be chosen to be close to 1, it follows

$$P(Y_i(\Delta s) > u) \to \mu(u)\Delta s \tag{44}$$

It is $P(Y_i \leq S) = P(\xi \leq S - g_i) = P(\xi \leq u_i)$ with $u_i = S - g_i$. In win (39) $P(\xi \leq u_i)$ may be replaced by (41); it follows

$$\begin{aligned} P(Y \leq S) &\approx \exp\left(\sum_{i=1}^{n} \log P(\xi \leq u_i)\right) \\ &\approx \exp\left(\sum_{i=1}^{n} (1 - (1 - \Phi(u_i) + \mu(u_i)\Delta s))\right) \end{aligned} \tag{45}$$

For $u_i \to \infty$,that is for $S \to \infty$ it follows that

$$P(Y \leq S) \approx \exp\left(\sum_{i=1}^{n} \log(\Phi(u_i) - \mu(u_i)\Delta s\right)$$

For $n \to \infty$ and therefore $\Delta s \to 0$ it follows, because of $\log(1-\epsilon) \approx -\epsilon$,

$$P(Y \leq S) \approx \exp\left(-\frac{\Delta x\sqrt{\lambda_2}}{2\pi}\sum_{i=1}^{n}\exp\left(-\frac{(S-g_i)^2}{2}\right)\right) \quad (46)$$

$$\approx \exp\left(-\frac{\sqrt{\lambda_2}}{2\pi}\int_{-a}^{a}\exp\left(-\frac{(S-g(x))^2}{2}\right)\,dx\right) \quad (47)$$

□

References

[1] Albert, P.S., McFarland, F.H., Smith, M.E., Frank, J.A. (1994) Time Series for modelling counts from a relapsing-remitting disease: application to modelling disease activity in multiple sclerosis. Statistics in Medicine, 13, 453-466

[2] Andersen, P.K., Borgan,O., Gill, R.D., Keiding, N.: Statistical models based on counting processes. Springer-Verlag, New York, Berlin, Heidelberg 1993

[3] Arnold, L. Stochastische Differentialgleichungen. R. Oldenburg Verlag München (1973)

[4] Bloxom, B. (1985a) A constrained spline estimator of a hazard function. Psychometrika, 50, 301-321

[5] Bloxom, B. (1985b) Considerations in psychometric modeling of response time. Psychometrika, 50, 383-397

[6] Cox, D.R. (1972) Regression models and life tables (with discussion). J. R. Stat. Soc. B. 34, 187-220

[7] Cox, D.R., Oakes, D. (1984) Analysis of survival data. Chapman and Hall, New York

[8] Cramer, H., Leadbetter, M.R. Stationary and related processes. John Wiley and Sons, New York 1967

[9] Ditlivsen, O. Extremes and first passage times with applications in civil engineering. Some approximative results in the theory of stochastic processes. Theses. Technical University of Denmark, Copenhagen 1971

[10] Fahrmeier, L., Tutz, G. Multivariate statistical modelling based on generalised linear models. Springer-Verlag, New York, Berlin, Heidelberg 1994

[11] Gardiner, C.W. Handbook of stochastic methods. Springer-Verlag, Berlin, Heidelberg, New York (1990)

[12] Greenhouse, J.B., Silliman, N.P. (1996), Applications of a mixture survival model with covariates to the analysis of a depression prevention trial. Statistics in Medicine, 15, 2077-2094

[13] Henderson, R. (1995) Problems and Prediction in survival-data analysis. Statistics in Medicine, 14, 161-184

[14] Horsthemke, W., Lefever, R. Noise-Iduced Transitions. Springer-Verlag Berlin Heidelberg New York 1984

[15] Kalbfleisch, J.D., Prentice, R.L. The statistical analysis of failure time data. John Wiley and sons, New York, 1980

[16] Leadbetter, M.R., Lindgren, G., Rootzen, H. Extremes and related properties of random sequences and processes. Springer-Verlag New York Heidelberg Berlin 1983

[17] Lindenberg, K., Shuler, K.E., Freeman, J. Lie, T.J. (1975) First passage time and extremum properties of Markov and independent processes. J. Stat. Phys., 12, 217-250

[18] Lindenberg, K., Sheshadri, V. (1979) Analytic theory off extrema I. Asymptotic theory for Fokker-Planck-processes. J.Chem. Phys. 71, 4075-4084

[19] Lindenberg, K., West, B.J. (1986) The First, the Biggest, and other such considerations. J. Stat. Phys. 42, 201-243

[20] Luce, R.D.: Response times - Their role in inferring elementary mental organisation. Oxford University Press, Clarendon Press, New York, Oxford 1986

[21] Mortensen, U. Suhl, U.

[22] Murray, J.D. (1989)Mathematical Biology. Springer-Verlag, Berlin

[23] Roufs, J.A.R, Blommaert, F.

[24] Weidlich, W., Haag, G. Concepts and Models of a Quantitative Sociology. The dynamics of interacting populations. Springer-Verlag Berlin Heidelberg New York 1983