THE "SICKLE-HYPOTHESIS" A TIME DEPENDENT POISSON MODEL WITH APPLICATIONS TO DEVIANT BEHAVIOR AND OCCUPATIONAL MOBILITY

Andreas DIEKMANN
Peter MITTER

Forschungsbericht/
Research Memorandum No. 171
June 1982

Contents

Zusammenfassung

In den Sozialwissenschaften liegen häufig Daten über die zeitliche Abfolge von Positionswechseln (event histories) vor, die durch einen zugrunde liegenden stochastischen Prozeß erklärbar sind. Beispiele sind abweichendes Verhalten, die Änderung der Berufsposition, geographische Mobilität, Ehescheidungen usf. In der Regel erweist sich die Annahme als realistisch, daß die Neigung zum Wechsel einer Position abhängig ist von der Verweildauer in der jeweiligen Position. Speziell bei den obigen Beispielen kann darüber hinaus vermutet werden, daß die (infinitesimale) Wahrscheinlichkeit eines Zustandswechsels mit zunehmender Verweildauer ansteigt, ein Maximum erreicht und dann wieder absinkt. Diese "Sichelhypothese" für den Zeitverlauf der Übergangsrate bildet den Kern unseres erweiterten Poisson-Modells. Da das Modell die Ableitung der Dichteverteilung der Ankunftszeiten gestattet, können die Parameter mittels der Maximum-Likelihood-Methode geschätzt werden, wenn die Häufigkeitsverteilung der Zeiten, zu denen ein Positionswechsel erfolgt, gegeben ist.

Das Modell wird auf Daten über abweichendes Verhalten und berufliche Mobilität angewandt, wobei die Sichelhypothese mit alternativen Erweiterungen des Poisson-Prozesses konfrontiert wird. Abschließend werden eine Reihe weiterer Modellimplikationen und Verallgemeinerungen präsentiert.

Abstract

Social scientists are frequently confronted with event history data which might be explained by a stochastic process. Examples are deviant behaviour, occupational or geographic mobility, marriage dissolution etc. Sometimes the assumption that the propensity to leave a position depends on the duration in that position seems to be realistic. In these examples one might expect that the rate of leaving a position initially increases, eventually reaches a maximum, and finally decreases again. The extended Poisson model studied in this papers exhibits such a pattern. If arrival time data is available, the parameters of the model can be estimated by maximum likelihood techniques. This model is applied to study deviant behaviour and occupational mobility. The results are compared with those from alternative extensions of the Poisson process. Finally further implications and extensions of this model are discussed.

Introduction:

Stochastic processes are very promising in modelling social behavior because they take in account two essential aspects: The dynamic aspect of social behavior and the fact that social processes are governed by probabilistic laws.

A well known model is the simple Poisson process and its counterpart, the exponential distribution of "arrival times". Many applications to social science data of the Poisson model and its extensions are reported in the literature (see for example COLEMAN 1964, chapt. 10 and 11). The simple Poisson process assumes a constant intensity, i.e. a constant probability that a new event will happen. But this model is not adequate if the probability of an event changes over time. More realistic are time dependent processes where intensity is a function of time.

An example of a time dependent Poisson process is the "Weibull-process" (MANN,SCHAEFER, SINGPURWALLA 1974, 127-129, see PETERSEN 1979 for an application). The model assumes the probability of an event to be a monotone function of time.

However, the Weibull-model cannot describe a process where the probability first increases and then - after reaching a maximum value - decreases asymptotically to zero. This type of a model where the transition rate as a function of time has the shape of an inverted U would be very useful in explaining a variety of social processes.

Let us take for example deviant behavior. There are good reasons to assume that the probability a person will move from the state of conformity to the state of deviance (i.e. the first deviant act) does depend on his age. For many delinquent acts the probability is low in early childhood, then increases till adolescence, and decreases thereafter.

This time path might apply to occupational mobility as well. There is a low likelihood of job shifts for people who have been in their positions for a short while and also for people who have been in their positions for a very long time. The maximal job shift probability is somewhere between these extremes.

Since the intensity rate as a function of time looks a sickle (inverted U) we call this process the "sickle-hypothesis". In this article we first construct an appropriate model for processes of that kind and then discuss the application of this model to data from criminology and occupational mobility.

I. The Model

We consider a model with two states. For the case of deviant behavior state 0 can be interpreted as the state of conformity and state 1 as the state of deviance. A change from 0 to 1 takes place if an individual commits the first delinquent act. State 1 is an absorbing state (there is no possibility to move back from state 1 to state 0).

Probability of change

We assume that for small time intervals the probability of change is asymptotically proportional to the length of the time interval:

(1) $$\text{Prob (change occurs in } [t,t+\Delta t]) = \alpha(t).\Delta t+o(\Delta t)$$

where $[t,t+\Delta t]$ denotes the time interval of length Δt starting at time t, and $o(\Delta t)$ is a function such that $\lim_{\Delta t \to 0} o(\Delta t)/\Delta t = 0$.

Notice that the process is not stationary (the intensity or hazard rate $\alpha(t)$ is a function of time). Thus the tendency to change the state varies with the age of the individuals under consideration.

Assumption (1) enables to compute the probability $p_o(t)$ of no change within time t:

(2) $$p_o(t+\Delta t) = p_o(t).[1-\alpha(t).\Delta t - o(\Delta t)] .$$

In the limit $\Delta t \to 0$ this becomes

(3) $$\frac{dp_o(t)}{dt} = -\alpha(t)p_o(t) .$$

This equation (3) states that the change in the probability of being in state 0 (the "outflow") is the negative of the product of the probability of being in state 0 times the tendency to move out of the state. The differential equation (3) can easily be solved for $p_o(t)$

(4) $$p_o(t) = \exp\left[-\int_0^t \alpha(\tau)d\tau\right]$$

Depending on $\alpha(t)$ the integral on the right hand side of (4) may or may not converge as time tends to infinity. In the ordinary Poisson process $\alpha(t)$ is a constant, so $\lim_{t\to\infty} p_o(t)=0$.

By contrast, in our model $p_I=\lim_{t\to\infty} p_o(t)$ can be positive. In this case p_I clearly denotes the probability that an individual will never change to state 1. This seems to be a very appealing feature of the model, but one should avoid interpreting p_I as a percentage of immune individuals. Indeed, since the fraction of the population which is affected by the deviance generating process is determined by the process itself, one cannot interpret the group of individuals who do not deviate as being unaffected by the process. In fact all individuals are exposed to the forces in favour of change, but it may happen that no change occurs. p_I is the corresponding probability.

Arrival times

The distribution function F(t) of the duration in the state of conformity (i.e. the arrival time T of a change) is simply the probability of a change within time t

(5) $$F(t) = 1-p_o(t) = 1-\exp\left[-\int_0^t \alpha(\tau)d\tau\right].$$

The density of the arrival time T is therefore

(6) $$f(t) = \alpha(t)\exp\left[-\int_0^t \alpha(\tau)d\tau\right].$$

Using (5) and (6) one obtains a nice expression for $\alpha(t)$

(7) $\alpha(t) = \frac{f(t)}{1-F(t)}$.

If the probability of no change at all p_I is positive, then T is an improper (or defective) random variable. Although this may cause some problems such as non-existence of a mean arrival time it does not affect the results in this paper.

Intensity

We assume that the intensity - thus the probability of change in an infinitesimal time interval - is determined by two competing factors

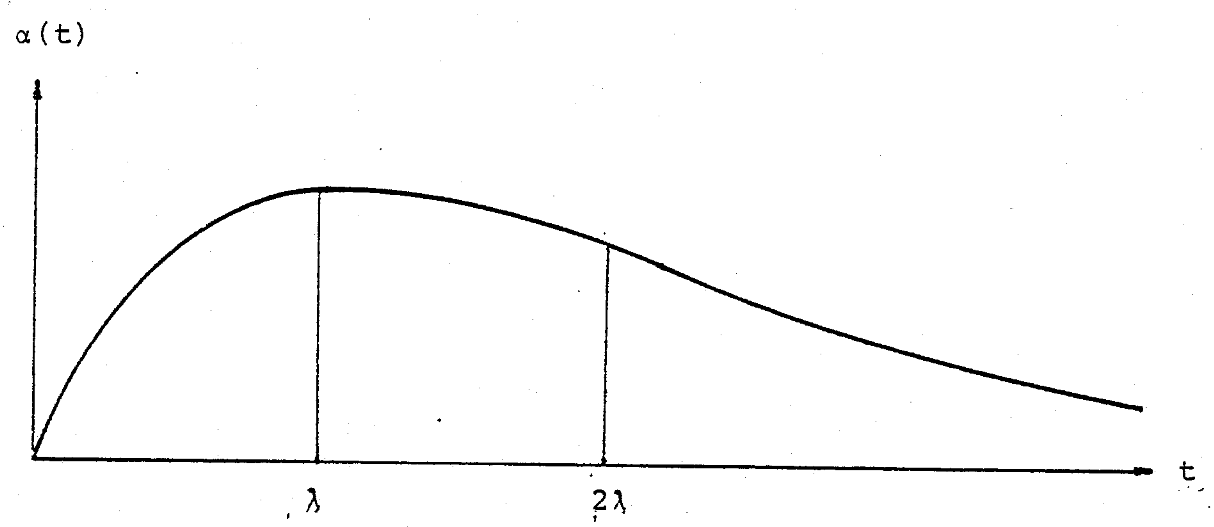
(i) a progressive factor which raises the probability of change with increasing age
(ii) a conservative, ultimately dominating factor which diminishes the probability of an occurence with increasing age.

The former corresponds to phenomena like imitation or increasing dissatisfaction (it is assumed to be proportional to duration t), the latter to increasing immunity because of risk aversion, expedicency or maturity (we assume this factor to decline exponentially). So we obtain the following form of the intensity function

(8) $\alpha(t) = c.t.e^{-t/\lambda} \quad c,\lambda>0 \quad t\geq 0$

Function (8) has some desired properties. At the first place it conforms properly with the sickle-hypothesis. It has one maximum point and one point of inflection and approaches zero in the limit. Secondly the function is an economic parametrization with two parameters which can be empirically meaningful interpreted. As can be seen from figure 1, λ is the time elapsed till the maximum point, the maximal intensity is $c.\lambda/e$. The point of inflection is located at $t=2\lambda$. If $\alpha(t)$ is multiplied by a suitable normalization factor it becomes a special case of the Γ-distribution. However, note that $\alpha(t)$ is not a density distribution but a hazard function.

Figure 1: The intensity function $c.t.e^{-t/\lambda}$



Distribution of arrival times

The probability of no change within time t is the solution of (4) with intensity (8)

(9) $$p_o(t) = \exp\{-\lambda c[\lambda-(t+\lambda)\exp(-t/\lambda)]\} ,$$

and has a positive limit $\exp(-\lambda^2/c)$, thus the random variable T (arrival time) is defective. Using (5), (6) and (9) the corresponding distribution and density function can be evaluated

(10) $$F(t) = 1 - \exp\{-\lambda c[\lambda-(t+\lambda)\exp(-t/\lambda)]\}$$

(11) $$f(t) = ct\exp(-t/\lambda)\ \exp\{-\lambda c[\lambda-(t+\lambda)\exp(-t/\lambda)]\}$$

The maximal density is located at the point $t_m<\lambda$ which is the solution of

(12) $$1 = t_m\left[1/\lambda + c\, t_m \exp(-t_m/\lambda)\right] .$$

Approximately,

(13) $$t_m \cong \frac{\lambda}{1+\lambda(c\lambda/e)}$$

II. Estimation of Parameters

Frequently data on arrival times are used to estimate the parameters of lifetime or failure models. This also seems to be appropriate in the present situation, but two problems arise.

The first problem concerns the measurement of the arrival time. In the case of criminal behavior the data might be the age of first offence. This age does not correspond to the concept of arrival time, because then the model would imply a high probability to behave illegally even for babies. Obviously it seems to be necessary to define a time origin with respect to age. This origin should reflect a stage of development of physical and mental ability to commit the offence under consideration. It may enter the model as a third parameter - say t_o - such that the intensity $\alpha(t)$ equals 0 for $t \leq t_o$ and $\alpha(t)=c(t-t_o) \exp(-(t-t_o)/\lambda)$ for $t \geq t_o$. Another way to deal with this problem would be to define this origin a priori due to additional considerations and then to calculate arrival time data before entering the estimation process. This is the way that threshold parameters are usually incorporated in lifetime models (see e.g. KALBFLEISCH and PRENTICE 1980 who argue that it would be rare that t_o would be known to exist without its value being known). We chose the smallest observed age at first offence as an a priori measure of t_o .[1)]

The second problem concerns the defectiveness of the arrival time variable. The maximum likelihood method which is used to estimate the process parameters is not defined for defective variables. To overcome this difficulty the model is modified in the following way: choose a priori a value T^* sufficiently high (higher than an individual's lifetime e.g.) and assume $\alpha(t)$ to be constant for all $t \geq T^*$. Then the resulting arrival time variable is no longer defective, and expressions (10) and (11) hold for $t \leq T^*$. As all observations will be within the range $[0,T^*]$, it is not necessary to know the numerical value of T^* . Although the process parameters are estimated for this modified model, their interpretation in terms of the original one is justified.

The maximum likelihood method (MML) chooses that pair $\hat{c}$ and $\hat{\lambda}$ among all possible pairs which make the observations most plausible, provided that the model is true. Technically it reduces to the problem of finding the maximizing arguments of the likelihood function $L(c,\lambda)$. This may be done by solving the first order conditions $\partial L/\partial c=0$, $\partial L/\partial \lambda=0$ or by an appropriate optimization technique. The likelihood function and the first order conditions are determined in the appendix. MML also yields estimates of the standard errors of the coefficients. These are usually needed to test whether individual coefficients are zero. This is not of importance in the present context, because $c=0$ or $\lambda=0$ implies that there are no cases of deviant behavior at all - a hypothesis obviously inconsistent with the observations.

III. Empirical Applications: Deviant Behavior and Occupational Mobility

We used DIEKMANN's (1980) survey on shoplifting[2] and an Austrian survey on occupational mobility[3] to test the sickle-hypothesis empirically. As can be seen from figures 2 and 3, the lifetable estimates[4] of the intensity functions are increasing at the beginning and decreasing in the later life . Four parametric models were estimated from the original arrival data[5]:

- Poisson process: $\alpha(t)$ = constant, exponential distribution of arrival times
- Weibull process: $\alpha(t) = \lambda p(\lambda t)^{p-1}$
- log-logistic process: $\alpha(t) = \lambda p(\lambda t)^{p-1}/(1+(\lambda t)^p)$, logistic distribution of logarithmic arrival times
- sickle curve: $\alpha(t) = c.t \exp(-t/\lambda)$

The estimated parameters are given in table 1, and the resulting arrival time distributions in tables 2 and 3. In both tables the last age group is considerably large, it contains censored times. The existence of a high percentage of censored observations with large values leads to unsatisfactory Poisson and Weibull estimates of arrival times for the part of the time scale covered in the tables. We thus consider the log-logistic and the sickle fit only in what follows. In both examples the log-logistic estimates fit slightly better than the sickle ones,but it is worth doing to study the differences in detail.

In the shoplifting study neither model yields satisfactory estimates of the intensity curve. The intensity is overestimated in the beginning and at later age and it is underestimated for the age when individuals are especially exposed to shoplifting for the first time. This age of maximal risk is overestimated, too. In contrast to the observations the

Table 1

ML-estimates of model parameters[6)]

Model	shoplift study λ	p	c	mobility study λ	p	c
Poisson	.0514			.0924		
Weibull	.0112	1.56		.0659	1.14	
log-logistic	.0762	2.12		.142	1.63	
Sickle	13.2		.0149	6.00		.0593

Table 2

Observed and fitted age at first shoplift

Age	Actual number of individuals	exponential fit	Weibull fit	log-logistic fit	sickle curve
4	1	12.0	.2	1.0	1.7
5	2	11.4	.4	3,3	4.7
6	6	10.8	.6	5.7	7.0
7	7	10.2	.7	7.8	8.9
8	8	9.8	.8	9.6	10.0
9	6	9.3	.9	10.8	10.8
10	14	8.8	1.0	11.7	11.2
11	8	8.4	1.0	12.1	11.3
12	26	7.9	1.1	12.1	11.1
13	19	7.5	1.2	11.9	10.8
14	15	7.1	1.2	11.4	10.3
15	9	6.8	1.3	10.8	9.8
16 or more *	118	129.0	228.9	130.7	131.3
total	239	239.0	239.0	239.0	239.0

* including individuals with no shoplift prior to interview (85)

expected maximal risk is fairly constant for the considerably broad age group from about 12 to 23 years. Figure 2 suggests that an adequate intensity curve should possess a second point of inflection prior to the age of maximal intensity. While this happens to be in the log-logistic model, it cannot be modelled by functions of type (8). So it seems that a more adequate parametrization of this process should have at least three parameters.

Corresponding to the flatness of the estimated sickle curve, the expected probability that a person will not deviate $p_I=0.074$ is quite small. 35,6 % of the individuals in the sample had not shoplifted prior to the interview, two-thirds of them were at least 20 years old. From these figures it seems that an estimate of roughly 0.2 for p_I is realistic. It should be noted that the exptected value of p_I in the log-logistic model is zero.

The fit of both models is much better in the mobility study. It seems that the sickle intensity declines too fast, but the corresponding probability of no change at all $p_I=.118$ sounds plausible (16 % of the individuals in the sample had no occupational shift prior to the interview, and all of them had been in their first position for at least 17 years). The overestimated intensity between the 5th and 12th year of carreer is longing for improvement nevertheless.

Table 3

Observed and fitted year of first shift in occupational position

year	Actual number of indivi- duals	exponential fit	Weibull fit	log-logistic fit	sickle curve
0	53	233.3	117.9	106.1	69.2
1	252	212.7	134.2	196.1	170.9
2	237	193.9	136.1	225.8	222.8
3	277	176.8	134.1	225.2	237.8
4	196	161.2	130.4	209.5	229.6
5	170	147.0	125.6	187.6	209.2
6	158	134.0	120.3	164.6	183.8
7	119	122.2	114.7	143.0	157.9
8	89	111.4	108.9	123.6	133.7
9	97	101.6	103.1	106.8	112.3
10	94	92.6	97.4	92.3	94.0
11	77	84.4	91.8	80.1	78.5
12	69	77.0	86.3	69.7	65.6
13	60	70.2	81.1	60.9	54.9
14	56	64.0	76.0	53.5	46.0
15	52	58.3	71.2	47.2	38.7
16	47	53.2	66.6	41.8	32.7
17 or more *	540	549.4	847.3	598.1	505.5
total	2643	2643.0	2643.0	2643.0	2643.0

* including individuals with no shift prior to interview (422)

Figure 2: Life table estimates and expected intensity curves of shoplifting data

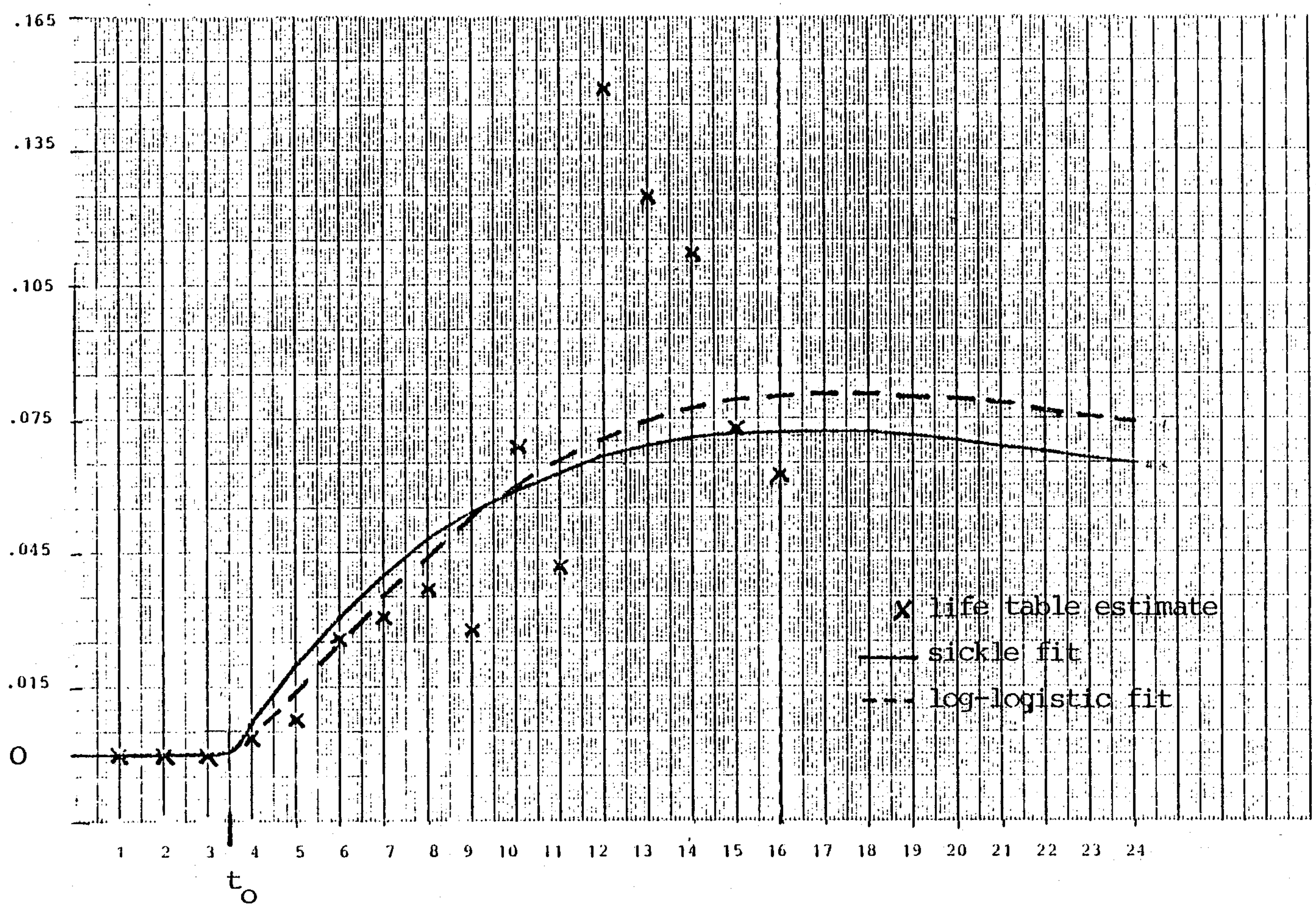
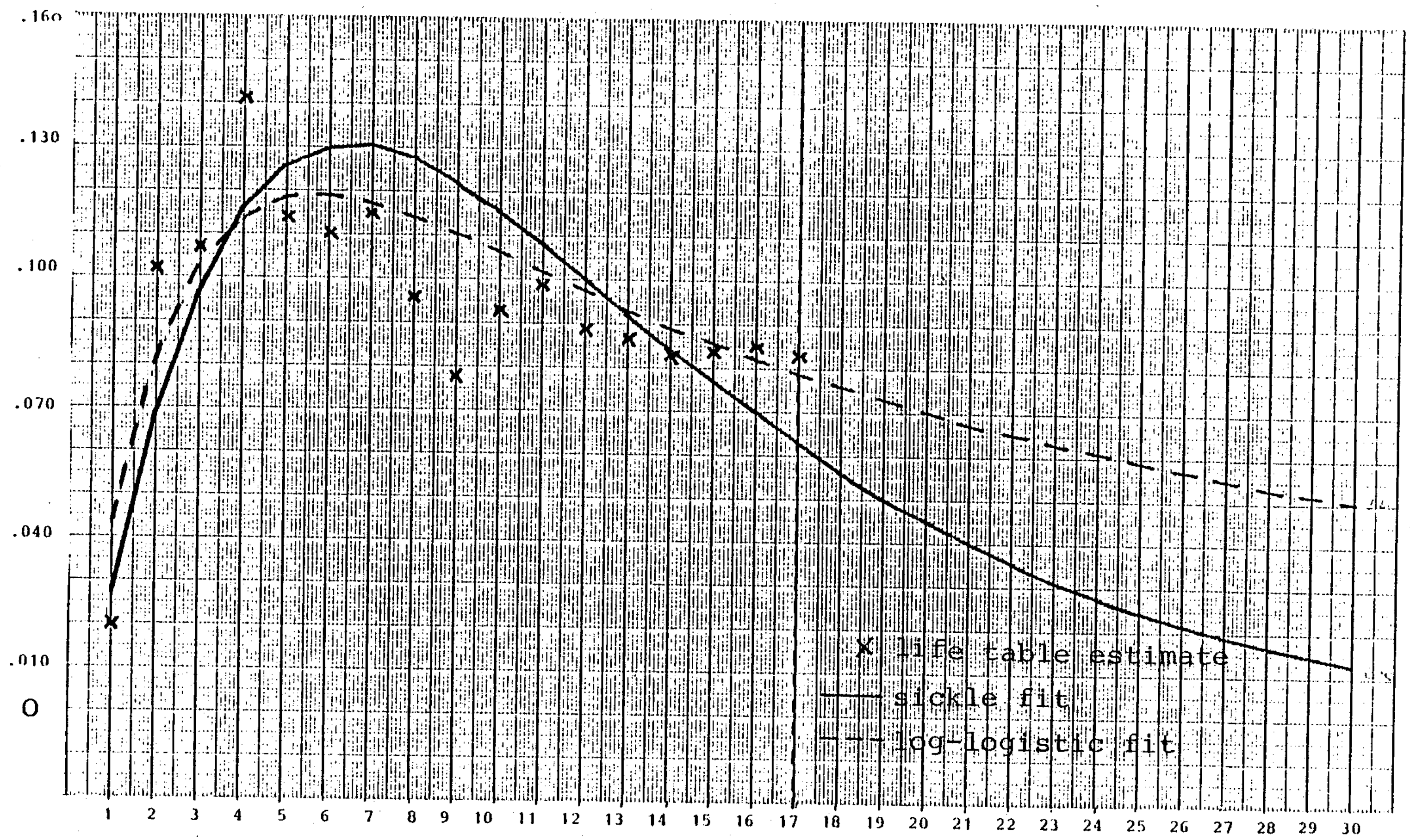


Figure 3: Life table estimate and expected intensity curves for occupational mobility data



IV. Extensions of the model

In this section we briefly mention possible lines of generalizations, further deductions from the model, and the relation to other stochastic processes.

Generalization of the "sickle": a bell shaped function

A time dependent intensity function with suitable properties can be obtained by adding a third parameter θ to the sickle function:

$$(14) \qquad \alpha(t) = c \cdot t^{\theta} e^{-t/\lambda}$$

$\alpha(t)$ takes it's maximum value at $t_{max}=\theta\lambda$ with inflection points located symmetrically at $t_{inf1,2}=t_{max}\pm\lambda\theta^{1/2}$.

Therefore, for $\theta>1$ the function is bell shaped with a second inflection point in the interval $\theta<t_{inf2}<t_{max}$. Our experience with actual data analysis leads to the conclusion that the presence of a second inflection point to the left side of the maximum is a particularly desirable property in view of the life table estimates (see figure 1 and 2). However, integration of $\alpha(t)$ yields an explicit solution only if θ is an integer value.

Event counts

So far interest has focused on the distribution of arrival times. If the assumption holds that intensity $\alpha(t)$ is independent of prior state shifts (i.e. the intensity to move from state k to k+1 is independent of k) the distribution of events happening in the interval $[0,t]$ follows the Poisson law (CHIANG 1968 : 49):

$$(15) \qquad p_k(t) = \frac{\exp\left\{-\int_0^t \alpha(\tau)\, d\tau\right\} \left[\int_0^t \alpha(\tau)\, d\tau\right]^k}{k!}$$

Of course, the assumption of no infection is debatable for the deviance data but it might be more appropriate for occupational shifts.

Random variable "number of events" $X(t)=k$ takes the expectation:

$$(16) \qquad E\left[X(t)\right] = \int_0^t \alpha(\tau)\, d\tau = \lambda c\left[\lambda-(t+\lambda)\ \exp\ (-t/\lambda)\right]$$

As can be seen from (16) the expected number of state shifts tends to the upper limit $c\lambda^2$ for time approaching infinity. If assumptions of the model are met (16) can be used to predict the average number of events for any chosen time interval $[0,t]$. The predicted behavior in time for occupational mobility and deviant behavior data is depicted in figure 4 (see table 4).

Figure 4: Expected number of events for deviant behavior and occupational mobility data

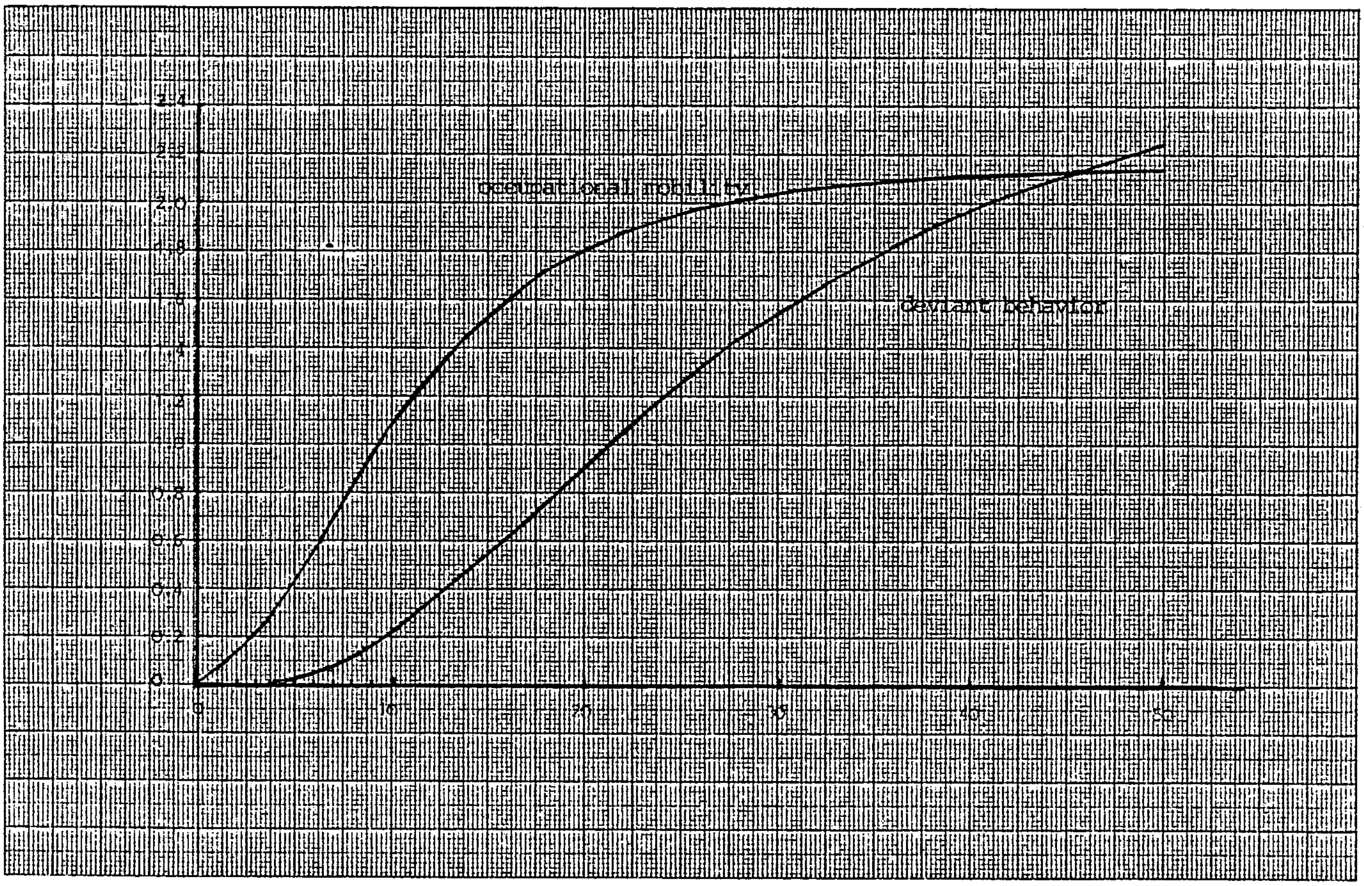


Table 4: Mean number of events in time span [o,t] for combination of sickle and Poisson model

t	Occupational mobility E[x(t)]	Deviant behavior E[x(t)]
0	0	0
1	.03	0
2	.10	0
3	.19	0
4	.31	.002
5	.43	.02
8	.82	.12
10	1.06	.23
15	1.52	.56
20	1.80	.92
30	2.05	1.55
40	2.11	1.98
50	2.13	2.25
⋮		
∞	2.13	2.60

Infection

In case of deviant behavior independence of intensity from prior deviant acts is not very plausible. In contrast positive contagion effects (criminal career hypothesis) or negative infection effects (deterrence hypothesis) might be expected. Although contagion is no problem if the model deals with the first delinquent act (shift from state 0 to state 1) as in sections I to III, this assumption is crucial for the distribution of event counts.[7)]

Using results of CHIANG (1980) a general contagion model was recently proposed by HAMILTON and HAMILTON (1981). Their basic idea is to model the intensity as a product of two functions: The first function expresses contagion, i.e. dependence of state k, while the second function represents the time dependency. If we assume sickle-type time dependency the intensity $\alpha_k(t)$ takes the following form:

$$(17) \qquad \alpha_k(t) = \alpha_k \cdot \alpha(t),$$

with $\alpha(t) = c\ t \exp(-t/\lambda)$ and α_k a function of k. The resulting distribution of event counts is a Chiang distribution as shown by HAMILTON and HAMILTON (1981).

For example, in the special case of linear contagion ($\alpha_k = A + Bk$) a modified version of CHIANG's (1964, chap.10) negative binomial distribution is obtained. The only difference to Coleman's distribution is the substitution of t by $\frac{1}{c}\int_0^t \alpha(\tau)d\tau$. This follows simply from the Chiang theorem presented in HAMILTON and HAMILTON (1981).

Heterogeneity

Allowing for heterogeneity is an extension in another direction. Assume the parameter c in the intensity function $\alpha(t)$ varies in the population according to a Γ-distribution.[8] Then, Poisson distribution (15) of event counts is conditional on c. The unconditional distribution is the following compound distribution (for details see CHIANG 1968: 49-50):

$$(18) \qquad p_k(t) = \int_0^\infty p_{(k/c)}(t) \cdot f(c) \cdot dc$$

with $p_{(k/c)}(t)$ identical to (15) and f(c) a Γ-distribution with parameters γ (argument of Γ-function) and β. The solution of integral (18) is a negative binomial distribution (CHIANG 1968):

$$(19)\qquad p_k(t) = \binom{k+\gamma-1}{k}\left(\frac{g(t)}{\beta+g(t)}\right)^k\left(\frac{\beta}{\beta+g(t)}\right)^\gamma$$

with $g(t) \equiv \int_0^t \tau e^{-\tau/\lambda} d\tau$.

From event count distribution (19) the arrival time distribution for the first occurrence can be easily derived:

$$(20)\qquad F(t) = 1 - p_0(t) = 1 - \left(\frac{\beta}{\beta+g(t)}\right)^\gamma$$

$$= 1 - \left\{\frac{\beta}{\beta+\left[\lambda^2-\lambda(t+\lambda)\ \exp\ (-t/\lambda)\right]}\right\}^\gamma$$

In the special case γ=1 this is the arrival time counterpart of a geometric distribution.

Finally we arrive at the density by differentiating (20) [9]:

$$(21)\qquad f(t) = \frac{\gamma\beta^\gamma\ t\ \exp[-t/\lambda]}{\{\lambda^2-\lambda(t+\lambda)\ \exp\ [-t/\lambda]+\beta\}^{(\gamma+1)}}$$

(21) is the p.d.f. of arrival times if heterogeneity is introduced in the suggested manner. That means if c is Γ-distributed in the population the sickle intensity function leads to (21) instead of (11).

<u>Cohort arrival counts</u>

Besides arrival times and event counts in time span [0,t] a third kind of distribution - connected with the same stochastic law - can be derived. Let $q_n(t)$=prob[Z(t)=n] , the probability that in time interval $[0,t]$ n units out of a cohort of N members will change state. The distribution follows from the differential equations of a "death process" (see LAND 1971 for an application) with intensity:

$$(22) \qquad \alpha_n(t) = \alpha(t)\cdot(N-n)$$

However, there is a very simple alternative way to find the distribution of Z(t). Remember that arrival time probability is $1-p_o(t)$ and survival probability is $p_o(t)$. Because occurrence of events are stochastically independent and because there are $\binom{N}{n}$ combinations of "arrivers" and "survivers" Z(t) follows a binomial distribution:

$$(23) \qquad q_n(t) = \binom{N}{n}\left[1-p_{o(t)}\right]^n \left[p_o(t)\right]^{(N-n)}$$

with expected value:

$$(24) \qquad E\left[Z(t)\right] = N\left[1-p_o(t)\right]$$

The mean number of shifts to state 1 is exactly the expected frequency distribution contained in tables 2 and 3 for the mobility and delinquency data.

Further applications

We suppose that the sickle model can be applied not only to criminal or occupational careers but to a variety of other processes. Let us illustrate this by a few examples.

For marriage and divorce patterns it seems to be very unlikely that people divorce immediately after wedding day or on the other hand after a golden wedding. However, there is a greater likelihood for divorce somewhere in between these events. This is at least true for Austrian and US-data on marriage cohorts (see FERRISS 1970, LAND 1971). Different parameters of the sickle model for Austria and the US would reflect cultural differences[10].

Two other examples are traffic accidents and problem solving in psychology of learning. The hazard rate for traffic accidents increases after people achieve their drivers' licenses. The reason is supposed to be overestimation of their driving abilities. Later on the subjective - objective ability gap is reduced either by more practice or by more realistic estimation of abilities. In contrast to a monotonous hazard model the sickle is expected to be a more adequate model for these data.

In psychology of learning, problem solving is often dependent on two time dependent factors, namely practice and motivation. If practice increases with time and the motivational factor decreases with time a sickle model with solution time as the random variable could be appropriate.

In general there are two arguments for the fruitfulness of a sickle type intensity function. From an empirical point of view life table estimates are not monotonic in many circumstances. Secondly from a more theoretical perspective substantial processes are very often governed by driving and inhibiting mechanisms, whereby ultimately the inhibiting force dominates. Therefore, the sickle model should be regarded as an alternative to monotonic increasing or decreasing probability laws. We think that it would be a useful avenue of research in the future to arrange tests of the sickle versus rival hypotheses with empirical data from different fields of application.

Acknowledgements

We are indebted to Elaine Bennett, Kenneth C. Land and Ian Patterson for helpful comments and to Gilg Seeber who estimated the parameters of the simple Poisson, Weibull, and log-logistic model using the program GLIM. This study has been supported by a grant of the "Deutsche Forschungsgemeinschaft" (DFG).

Appendix: Maximum Likelihood estimation of parameters

Given a set S of mutually independent observations, define S_o to be the subset of cases where no deviance has been observed, and S_1 to be the complementary subset where deviance has been observed. For the observations in S_1 let t_i be the observed arrival time of the change to deviance, for S_o let t_i be the time elasped between the start of the process and the censoring event (e.g. the interview). Then the likelihood function is given by

$$(25) \qquad F^*(c,\lambda) = \prod_{S_1} f(t_i) \prod_{S_o} (1-F(t_i))$$

$$= \prod_{S} \exp\left[-\int_o^{t_i} \alpha(\tau)d\tau\right] \prod_{S_1} \alpha(t_i) =$$

$$= \prod_{S} \exp\{-c\lambda\left[\lambda-(t_i+\lambda)\ \exp(-t_i/\lambda)\right]\}\prod_{S_1} c\ t_i \exp(-t_i/\lambda)$$

The log likelihood function is

$$(26) \qquad F(c,\lambda) = \sum_{S} -c\lambda\left[\lambda-(t_i+\lambda)\ \exp(-t_i/\lambda)\right] + |S_1|\log c +$$

$$+ \sum_{S_1} \log t_i - \sum_{S_1} \frac{t_i}{\lambda}$$

where $|S_1|$ is the number of cases where deviant behavior was observed. The maximum likelihood estimates $\hat{c},\hat{\lambda}$ are those values which make $F(c,\lambda)$ - or, equivalently, $F^*(c,\lambda)$ - as large as possible.Differentiation with respect to c and λ yields the first order conditions

$$\frac{\partial F}{\partial c} = - \sum_{S} \lambda[\lambda - e^{-t_i/\lambda}(t_i+\lambda)] + \frac{1}{c}|S_1| = 0$$

$$\frac{\partial F}{\partial \lambda} = -c\sum_{S} 2\lambda - e^{-t_i/\lambda}[t_i+2\lambda+\frac{t_i(t_i+\lambda)}{\lambda}] + \sum_{S_1} \frac{t_i}{\lambda 2} = 0$$

The first condition implies

$$(27) \qquad c = \frac{|S_1|}{\sum_{S} \lambda[\lambda - e^{-t_i/\lambda}(t_i+\lambda)]}$$

After substitution of this expression into $F(c,\lambda)$ the problem is reduced to evaluating the maximizing value $\hat{\lambda}$ of

$$F(c,\lambda) = \text{const} - |S_1|\log(\sum_{S}\lambda[\lambda - e^{-t_i/\lambda}(t_i+\lambda)]) - \frac{1}{\lambda}\sum_{S_1} t_i$$

(notice that F is a function of only one variable λ). Subsequently, $\hat{\lambda}$ is used to calculate $\hat{c}$ by means of (27).

Notes

1) As can be seen from the frequency distribution contained in table 2 one respondent reported age 4 for the first event. Therefore the exact time point must be somewhat in between the interval from 3.5 to 4.5 . We chose the lower bound of the interval as the starting time, i.e. $t_o = 3.5$.

2) Survey data were collected by retrospective questionnaires administered to 241 German apprentices and college students in 1979.

3) Subsample (N=2643 men) from a retrospective survey on occupational carreers, conducted by the Austrian Bureau of Census in 1972.

4) For the life-table estimator formulae, see e.g. Kalbfleisch and Prentice 1980, p. 16.

5) The Poisson, Weibull and log-logistic estimates were accomplished by Gilg Seeber, using GLIM.

6) The parameters λ and p have a distinct meaning in different models, so they are not comparable.

7) In principle there are two types of contagion: "intra-career contagion" and contagion between persons. Only the first type of infection is no problem if shifts from state zero to state one are considered. If there is contagion of the second type the basic assumption of independence of events does not longer hold.

8) There are two reasons that heterogeneity is often introduced by a Γ-distribution. At the first place the Γ-distribution is very general containing other distributions like exponential or Chi-Square as special cases. Secondly the

model is mathematically tractable by analytic tools.

9) Alternatively the compound distribution (21) can be derived directly from distribution (11) with c following the Γ -distribution. Note that random variable T is defective.

10) At present the authors are conducting a project focusing the analysis of divorce data by means of stochastic models.

References

Chiang, Chin Long (1968), Introduction to Stochastic Processes in Biostatistics. New York: Wiley

Chiang, Chin Long (1980), An Introduction to Stochastic Processes and Their Applications. New York: Krieger

Coleman, James S. (1964), Introduction to Mathematical Sociology. New York: Free Press

Diekmann, Andreas (1980), Die Rolle von Normen, Bezugsgruppen und Sanktionen bei Ladendiebstählen. Research Memorandum No 156 of the Institute for Advanced Studies, Vienna. Vienna: Verlag "Wissenschaft und Technik"

Ferris, Abbott L. (1970, An indicator of marriage dissolution by marriage cohort. Social Forces 48: 356-365.

Hamilton, James D. and Lawrence C. Hamilton (1981), Models of social contagion. Journal of Mathematical Sociology 8: 133 - 160

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

Land, Kenneth C. (1971), Some exhaustible Poisson process models of divorce by marriage cohort. Journal of Mathematical Sociology 1: 213 - 232

Mann, N.R., R.E. Schaefer and N.D. Singpurwalla (1974), Methods for Statistical Analysis of Reliability and Life Data. New York: Wiley

Petersen, Ib Damgaard (1979), Laws of development of international systems. Methoden en Data Nievwsbrief van de Sociaal Wetenschappelijke Sectie van de Vereniging voor Statistiek (of Netherland) 4: 3 - 18