Categorical time series in psychological measurement
van Rijn, P.W.

Citation for published version (APA):
4

LOGISTIC MODELS FOR SINGLE-SUBJECT TIME SERIES

4.1 Introduction

Statistical methods in psychology are mostly applied to a collection of individuals rather than to a single one (Kratochwill, 1978, p. 3). The development of methods for psychological testing in the first half of the twentieth century put the individual on the background, because the initial objective of psychological testing was to differentiate among individuals. Keeping this in mind, the focus in the remainder of the twentieth century on advancement of statistical methods based on variation between individuals (inter-individual variation, IEV) instead of variation within a single individual, seems understandable. However, models for time-dependent variation of a single individual (intra-individual variation, IAV) have been widely available for some time. The discovery of the intrinsically stochastic time-dependent behavior within grains of pollen suspended in air (Brownian motion) led to the development of appropriate statistical models for single systems in the beginning of the 20th century. In this regard, the lack of interest in a pure $N = 1$ perspective in psychometrics seems remarkable.

It is not to say that examples of analyses of IAV are wholly absent in the psychometric literature. The measurement of (individual) change, for example, is a branch of psychometrics with a relatively long history. An early overview of problems encountered in measuring change can be found in Harris (1963). In that book, a single-subject analysis of multivariate time series is described by Holtzmann, who stressed that psychologists should study this type of analysis, in

---

view of the increasing importance of time series in other branches of science such as econometrics and biometrics (Holtzmann, 1963, p. 199). More recently, Nesselroade and Schmidt McCollam (2000) advocated analysing IAV in the context of developmental processes in psychology, and Collins and Sayer (2001) provided an overview of newly developed methods for the analysis of change.

Apart from the historical development, it is difficult to find an explicit and convincing rationale for the one-sided focus on IEV in contemporary psychometrics. The restriction to IEV appears to be considered to be an almost self-evident consequence of the scientific ideal to strive for general nomothetic knowledge. The science of psychology should involve theories and laws that apply to all human subjects. Such nomothetic knowledge would seem to be poorly served by intensive study of single subjects, because results thus obtained may not be generalizable in the intended sense. Despite its possible appeal, we will argue that this kind of rationale is incorrect in many instances by making use of a set of well-known mathematical theorems.

In our criticism of the one-sided focus on IEV, we do not take issue with the ideal of nomothetic knowledge, i.e., the search for psychological theories and laws that apply to all human subjects. Our criticism only concerns the assumption that theories and laws based on analysis of IEV apply to each human subject, and thus, would hold for IAV. To obtain valid theories and laws about IAV, one cannot generalize results derived from IEV, but one has to study IAV in its own right. That is the implication of the mathematical theorems to which we will refer. Having available the results of a sufficient number of individual analyses of IAV, one then can search for general characteristics by means of standard inductive techniques. If successful, this will yield valid nomothetic knowledge about the structure of IAV, i.e., nomothetic theories and laws about idiographic (individual) processes.

This chapter is divided into two parts. The first part starts out with a description of analyses for IEV and IAV, and the condition under which there exists a relationship between the two types of analysis, namely ergodicity. This condition is explained in the context of psychometrics.

Having thus set the stage for serious consideration of IAV, the second part of this chapter discusses latent variable models for single-subject time series data. Special attention is given to the logistic model for multivariate dichotomous time series, which can be seen, in its simplest form, as a dynamic variant of the Rasch model, where the person parameter is replaced by a person process. It must be stated that logistic models for repeated measurements have been discussed
4.2 The ergodic notion in psychology

In its basic form, standard statistical analysis in psychology proceeds by drawing a sample of individuals, assessing their scores on selected measurement instruments, and then computing statistics by taking appropriate averages over the scores of all available individuals. If all individuals would yield the same score, statistical analysis would be severely reduced. Hence, it is the manner in which scores vary across subjects, IEV, which provides the information for the analysis. In contrast, in time series analysis the same individual subject is repeatedly measured, and statistics are computed by taking appropriate averages over the scores obtained at all measurement occasions. Hence, it is the manner in which an individual’s scores vary across measurement occasions, IAV, which provides the information for time series analysis.

We already indicated that psychometricians are mainly interested in analyses of IEV. A vivid illustration of this tendency can be found in the classic treatise of test theory by Lord and Novick (1968). They define the concept of true score of a person as the mean of the distribution of scores obtained by independent repeated measurement of this person. This is obviously a definition in terms of IAV. Lord and Novick then remark that repeated measurement of the same person will affect this person’s state and give rise to fatigue, habituation, or other confounding effects. They conclude that therefore, instead of measuring one person a large number of times, test theory has to be based on the alternative paradigm in which a large number of persons is measured once or twice. The shift to the latter alternative paradigm implies that test theory is based on analysis of IEV.

Notwithstanding that confounding factors such as habituation and fatigue might complicate the implementation of a purely IAV based test theory, a reference to such contingent states of affairs cannot be taken as the reason for the impossibility of this whole paradigm. In addition, Lord and Novick (1968, p. 32) state that the definition of true score in terms of IAV would be better suited for individual assessment than an IEV based test theory that is meant to differen-
It might therefore be expected that psychological tests constructed on the basis of analysis of IEV perform suboptimally when applied for the purpose of individual prediction. However, a task that still awaits further elaboration is the assessment of situations in which such test performance is suboptimal.

The urgency to determine the performance of standard tests in the context of individual assessment and prediction is all the more pressing given the strong justification for the conjecture that the differences between analysis of IEV and IAV go deeper than a mere difference in degree of success in the context of individual prediction. The reasons we have in mind are of two kinds: the implications of ergodic theorems, and results from mathematical biology suggesting the presence of substantial heterogeneity in human populations. Ergodic theory concerns the characterization of stochastic processes for which analysis of IEV and IAV yield the same results (e.g., Petersen, 1983). The classic ergodic hypothesis originates from statistical physics, and states that the average of a stochastic process over time is equal to the average of the ensemble of stochastic processes at a single point in time. By an ensemble is meant the possibly infinite number of hypothesized copies of a system. An ergodic hypothesis can also be stated for the variance or distribution of a stochastic process. In this sense, for ergodic processes the one-sided psychometric focus on analysis of IEV does not present any problem, because results thus obtained also are valid for individual assessment and prediction of IAV. Unfortunately, however, the criteria for ergodicity are very strict, and involve the absence of any time-dependent changes in the distributional characteristics of a stochastic process. Therefore, all developmental, learning and adaptive processes do not obey the criteria for ergodicity. For these classes of non-ergodic processes, there may not exist any lawful relationship between IEV and IAV.

Related to ergodicity is the notion of stationarity, which concerns the distributional characteristics of a single realization of a stochastic process. Stationarity amounts to the absence of time-dependent changes in distributional characteristics and, except for certain special cases of non-stationarity, is a condition for ergodicity. For Gaussian processes, stationarity is a necessary condition for ergodicity. An example of a stationary process can be given by the notion of general intelligence in normal adults. Now, it can be assumed that, under nor-

\[\text{Note that strict stationarity is meant here, see, e.g., Hamilton (1994, p. 45-46).}\]

\[\text{For the sake of argument, let us neglect all problems associated with the theoretical status and operationalization of general intelligence.}\]
mal circumstances, a normal adult’s level of general intelligence does not change structurally over his lifespan. Administration of several intelligence tests over the lifespan will result in scores that vary, but probably only slightly. The distribution of intelligence scores in the first half of the measurements is likely to be equal to that of the second half, and the process can be considered to be stationary. However, if circumstances change drastically due to, for example, illness or excessive training on intelligence tests, this distribution is bound to change as well. The process then is no longer stationary.

Even if the distributional characteristics of a stochastic process are invariant in time, that is, the process is stationary, it still may be non-ergodic. The key difference between stationarity and ergodicity concerns the uniqueness of the so-called equilibrium distribution of a stochastic process, i.e., the distribution of the values of a stochastic process as time increases without bound. Each stationary process gives rise to an equilibrium distribution, but this equilibrium distribution may not be unique. Only if the process is ergodic, then this is necessary and sufficient for its equilibrium distribution to be unique (cf. Mackey, 1992, Theorem 4.6). Hence stationary processes are non-ergodic if they display a moderate kind of heterogeneity: their equilibrium distribution is not unique. Notice that this is the kind of non-ergodicity known from Markov chain theory (e.g., Kemeny, Snell, & Knapp, 1966). Already the presence of this moderate form of heterogeneity with respect to the equilibrium distribution implies the possibility of a lack of lawful relationships between IEV and IAV.

Now, let us return to our intelligence example. The general level of intelligence of an ensemble of normal adults is not likely to change structurally over time under normal circumstances. Yet, it is unlikely that all adults have the same distribution of intelligence scores over time, that is, there does not exist a unique equilibrium distribution. Thus, the ensemble of human adults in this case is non-ergodic, although the individual intelligence processes in this example can be considered stationary. If the ensemble of adults would be ergodic, the following odd statement would hold: "Five percent of the people score 125 or higher on an intelligence test, therefore five percent of the time your intelligence score is higher than 125".

There are strong indications that heterogeneity in human populations may be much more pervasive, transcending the moderate forms associated with non-ergodicity. Mathematical theory about biological pattern formation (e.g., Murray, 1993) and nonlinear epigenetics (Edelman, 1987) shows that growth processes are severely underdetermined by genetic and environmental influences. Conse-
sequently, growth processes have to be self-organizing in order to accomplish their
tasks. In particular the maturation of the central nervous system results from
self-organizing epigenetic processes. Self-organization, however, gives rise to sub-
stantial endogeneous variation that is independent from genetic and environmen-
tal influences (Molenaar, Boomsma, & Dolan, 1993; Molenaar & Raijmakers,
1999). For instance, homologous neural structures on the left-hand and right-
hand side of the same individual (IAV) can differ as much as the left-hand side
of this neural structure in different individuals (IEV). Insofar as the activity of
such heterogeneous neural structures is associated with the performance on psy-
chological tests, this performance can be expected to be heterogeneous in much
stronger forms than is the case with non-ergodicity.

It has been shown by means of simulation experiments as well as mathemat-
ical proof (Molenaar, Huizenga, & Nesselroade, 2003; Kelderman & Molenaar,
2007) that standard factor analysis of IEV is insensitive to the presence of sub-
stantial heterogeneity. For instance, it is an assumption of the standard factor
model that factor loadings are invariant (fixed) across subjects. If, however,
these factor loadings in reality varied randomly across subjects (a violation of
the assumption of fixed factor loadings), then the standard factor model still fits
satisfactorily. There appears to be only one principled way in which the presence
of such heterogeneity can be detected, namely by carrying out replicated factor
analyses of IAV (dynamic factor analysis of multivariate time series; Molenaar,
1985) and then compare the solutions thus obtained for distinct subjects.

In closing this section, it is reiterated that in general one cannot expect law-
ful relationships to exist between the structure of IEV and the structure of IAV.
Such relationships can only be obtained under the restrictive condition that the
processes concerned are ergodic. For non-ergodic processes, and in cases where
human subjects are heterogeneous in even more pervasive ways (e.g., each subject
having its personal factor model with its own distinct number of factors, factor
loading pattern and/or specific variances), the use of IAV paradigms is manda-
tory. To accomplish this, appropriate time series analysis extensions of standard
statistical techniques are required. Brillinger (1975) presents a rigorous deriva-
tion of time series analogues of all standard multivariate techniques (analysis of
variance, regression analysis, principal component analysis, canonical correlation
analysis). In the next section, we present an overview of time series analogues of
latent variable models.
4.3 Latent variable models

From a general point of view, a stochastic process can be interpreted as a random function. That is, as an ensemble of time-dependent functions on which a probability measure is defined (e.g., Brillinger, 1975, section 2.11). Each time-dependent function of this ensemble is called a trajectory (or realization). Even if information is available about the entire past of a stochastic process up to some time $t$, then exact prediction for the next time point still is impossible. Each trajectory in an ensemble extends over the entire time axis. An observed time series, i.e., the particular stretch of values obtained by repeated measurement of a single subject, constitutes a randomly drawn trajectory from the ensemble, where this trajectory is clipped by a time window with width equal to the period of repeated measurement. In what follows we will denote a stochastic process by $y_t$ and an observed time series thereof by $y_t$, $t = 1, \ldots, T$. We acknowledge that this notation is not entirely correct, but it is convenient and customary.

A subset of latent variable models for IAV is obtained by replacing all random variables in a standard latent variable model by stochastic processes. Bartholomew (1987) has given a useful classification of standard latent variable models based on two features: whether the observed variable is continuous or discrete and whether the common latent variable is continuous or discrete. This classification will be followed in our overview of latent variable models for IAV. There is an additional third feature which has to be considered for latent variable models for IAV, namely whether the time dimension is continuous or discrete. We will, however, restrict attention to models in discrete time only, as this is sufficient for our present purposes.

If both the observed variable and the common latent variable are continuous, the latent variable model is classified as a factor model. Replacement of all random variables in the linear factor model by continuous stochastic processes yields the linear state-space model: $y_t = Z_t \alpha_t + \epsilon_t$, wherein $y_t$ is the observed continuous $n$-variate process, $Z_t$ is a matrix of factor loadings, $\alpha_t$ is a common $m$-variate latent process (also called state process), and $\epsilon_t$ is a $n$-variate measurement error process. Statistical analysis of the state-space model is well developed and is treated in several text books (e.g., Durbin & Koopman, 2001). Hamaker, Dolan, and Molenaar (2003) discuss applications of the state-space model in psychological research.\footnote{Software for the fit of state-space models can be downloaded from: http://users.fmg.uva.nl/cdolan/}
If both the observed variable and the common latent variable are discrete, the latent variable model is classified as a latent class model. Replacement of all random variables in the latent class model by discrete stochastic processes yields the hidden Markov model (e.g., Elliott, Aggoun, & Moore, 1995). Visser, Raijmakers, and Molenaar (2000) present applications of hidden Markov modelling in psychological research.\(^5\) If the observed variable is continuous and the common latent variable is discrete, the latent variable model is classified as a latent profile model (Bartholomew, 1987; Molenaar & von Eye, 1994). Replacement of the observed variable by a continuous stochastic process and the common latent variable by a discrete stochastic process yields a variant of the hidden Markov model (Elliott et al., 1995).

If the observed variable is discrete and the common latent variable is continuous, the latent variable model is classified as a generalized linear model. Replacement of all random variables in the generalized linear model by discrete (observed) and continuous (latent) stochastic processes yields a dynamic generalized linear model as described in Fahrmeir and Tutz (2001).

We will focus on a subset of dynamic generalized linear models. That is, models in which the observed process is dichotomous, related to a continuous latent process through the logistic response function.

### 4.4 A LOGISTIC MODEL FOR DICHOTOMOUS TIME SERIES

Dichotomous (or binary) time series can be modelled in various ways. If auxiliary information is available, regression models can be used. For dichotomous time series, such models are discussed in detail in Kedem and Fokianos (2002) and in Fahrmeir and Tutz (2001). Our focus is on modelling dichotomous time series using latent variables which is comparable to the modelling of dichotomous variables in item response theory (see also, Mellenbergh, 1994; Mellenbergh & Van den Brink, 1998). Since the latent variable is replaced by a stochastic process, this approach can be seen as a dynamic extension of item response modelling. As stated before, this approach is not new, although the emphasis on the modelling of IAV in this sense is novel. Modelling is pursued following the dynamic generalized linear modelling approach as described in Fahrmeir and Tutz (2001), that is, by specifying a distributional model, response function, linear predictor, and transitional model.

\(^5\)Appropriate software can be found at: http://users.fmg.uva.nl/ivisser/hmm.
4.4 A logistic model for dichotomous time series

4.4.1 General outline

Consider the situation in which we have a dichotomously scored, multivariate time series, i.e., an \( n \)-dimensional observation vector \( y_t \) such that \( y_t \in \{0, 1\}^n \), at each time point \( t = 1, \ldots, T \). Each single univariate observation \( y_{it}, i = 1, \ldots, n \), follows a Bernoulli distribution with parameter \( \pi_{it} \) as the probability of obtaining a score one, given by

\[
y_{it} \sim B(\pi_{it}) = \pi_{it}^{y_{it}} (1 - \pi_{it})^{1 - y_{it}}, \quad 0 < \pi_{it} < 1.
\]  

(4.1)

The probability \( \pi_{it} \) is modelled by inserting a linear predictor \( \eta_{it} \) into the logistic response function, resulting in

\[
\pi_{it} = \frac{\exp(\eta_{it})}{1 + \exp(\eta_{it})}.
\]  

(4.2)

Next, the \( n \)-dimensional linear prediction vector \( \eta_t \) is constructed by linking the \( n \times m \) design matrix \( Z_t \) with the \( m \)-dimensional latent state vector \( \alpha_t \)

\[
\eta_t = Z_t \alpha_t.
\]  

(4.3)

A linear transition equation is assumed, which relates states at \( t - 1 \) to \( t \) through the \( m \times m \) transition matrix \( F_t \), and is given by

\[
\alpha_t = F_t \alpha_{t-1} + R_t \xi_t, \quad \xi_t \sim N(0, Q_t).
\]  

(4.4)

The state vector \( \alpha_t \) is allowed to contain time-invariant elements. The \( m \times p \) selection matrix \( R_t \) is assumed to be a subset of the columns of the \( m \)-dimensional identity matrix \( I_m \), so that it associates the elements of the \( p \)-dimensional disturbance vector \( \xi_t \) with the \( p \) time-varying elements of the state vector (Durbin & Koopman, 2001, p. 38).\(^6\) The elements of \( \xi_t \) are generally referred to as innovations. In general, the state process is started up by an initial state \( \alpha_0 \). The initialization is dependent on the type of process that is used. If all elements of the state vector are time-varying, the process can be initialized by \( \alpha_0 \), which is normally distributed as follows

\[
\alpha_0 \sim N(a_0, Q_0).
\]  

(4.5)

In the above representation, the covariance matrices \( Q_t \) are nonsingular, which is somewhat more advantageous in estimation procedures (Durbin & Koopman,

\(^6\)Notice that \( m = p + n \) does not necessarily follow.
2001, p. 38). Note that $a_0$, $Q_0$, $Z_t$, $F_t$, $R_t$, and $Q_t$ can contain parameters to be estimated. Methods for the estimation of these parameters are not well developed. However, Fahrmeir and Wagenpfeil (1997) describe an estimation procedure for the estimation of $a_0$, $Q_0$, and $Q_t$.

The following three assumptions are stated to completely specify the model in terms of densities. The first assumption is that current observations are dependent on current states only:

$$p(y_t|\alpha_t, \alpha_{t-1}, \ldots, \alpha_0, y_{t-1}, y_{t-2}, \ldots, y_1) = p(y_t|\alpha_t, y_{t-1}, y_{t-2}, \ldots, y_1).$$

The second assumption is that the state process is first order Markovian:

$$p(\alpha_t|\alpha_{t-1}, \ldots, \alpha_0) = p(\alpha_t|\alpha_{t-1}).$$

Finally, and in addition to assumption one, it is assumed that the multivariate observations are independent given the current state:

$$p(y_t|\alpha_t, y_{t-1}, y_{t-2}, \ldots, y_1) = \prod_{i=1}^{n} p(y_{it}|\alpha_t, y_{t-1}, y_{t-2}, \ldots, y_1).$$

Since the specific contents of the state and disturbance vector can be freely chosen, a variety of latent processes can be captured with the current representation. Depending on the hypothesized dynamic constellation of the latent process, one can choose between, for instance, autoregressive processes, moving average processes, and random walks (e.g., Hamilton, 1994). In addition, trends and cyclic change parameters can be included in the current representation. For now, we restrict the latent process to be either a white noise process, an autoregressive process, or a random walk.

The model can be extended to more than one person ($N > 1$), more than one latent process ($p > 1$), and also to multi-categorical or polytomous time series. However, the interest here lies in $N = 1$ and since results of analyses of dichotomous time series with this type of models are not widely available, we next consider some simple, yet illustrative modelling examples.

### 4.4.2 A Dynamic Logistic Model

We now illustrate how a dynamic variant of the Rasch model can be obtained. We begin by constructing the state vector $\alpha_t$, which consists of two parts. The first part describes a person’s univariate latent process $\theta_t$. The second part consists of $n$ time-invariant threshold parameters, each associated with the corresponding
element of \( y_t \), denoted by \( \beta \). We consider three different processes, namely, a white noise process, a first-order autoregressive process, and a first-order random walk. These processes are given by, respectively,

\[
\begin{align*}
\theta_t &= \mu_\theta + \xi_t, & \xi_t &\sim N(0, q), \\
\theta_t &= \mu_\theta + \phi_1 \theta_{t-1} + \xi_t, & \xi_t &\sim N(0, q), \\
\theta_t &= \mu_\theta + \theta_{t-1} + \xi_t, & \xi_t &\sim N(0, q),
\end{align*}
\]

where \( \mu_\theta \) is a time-invariant mean and \( \phi_1 \) is the autoregressive parameter. Note that if \( |\phi_1| < 1 \), the autoregressive process is stationary. The random walk in Equation 4.8 can be perceived of as the discrete time analogue of Brownian motion (Klebaner, 1998, p. 80). It should be noted that the random walk process is nonstationary since \( \text{Var}(\theta_t) \rightarrow \infty \) as \( t \rightarrow \infty \), and therefore non-ergodic. For each of the three processes, we have \( \alpha_t = (\theta_t, \mu_\theta, \beta')' \) and \( m = p + n + 1 \). For simplicity and sufficiency for present purposes, the following model parameters and matrices are considered time invariant as well: the design matrix \((Z)\), the transition matrix \((F)\), the selection matrix \((R)\), and the covariance matrix of the state disturbances \((Q)\).

Consider now the situation in which we have four dichotomous variables. Modelling is pursued as follows. We have a 4-dimensional vector of observations \( y_t \), a 4-dimensional probability vector \( \pi_t \), and a 4-dimensional linear prediction vector \( \eta_t \), related to each other as stated in Equations 4.1 and 4.2. The time invariant elements of the state vector do not need to be initialized, and the white noise process in Equation 4.6 does not either. Only the autoregressive process and the random walk have to be initialized with \( \theta_0 \). The initial state and the 6-dimensional state vector then have the following form

\[
\alpha_0 = R\theta_0 = \begin{bmatrix}
\theta_0 \\
0 \\
0 \\
0 \\
0 \\
0
\end{bmatrix}, \quad \alpha_t = \begin{bmatrix}
\theta_t \\
\mu_\theta \\
\beta_1 \\
\beta_2 \\
\beta_3 \\
\beta_4
\end{bmatrix}.
\]

The specification of the design matrix defines the relation between the person
process and the threshold parameters, and is given by

\[
Z = \begin{bmatrix}
1 & 1 & -1 & 0 & 0 & 0 \\
1 & 1 & 0 & -1 & 0 & 0 \\
1 & 1 & 0 & 0 & -1 & 0 \\
1 & 1 & 0 & 0 & 0 & -1 \\
\end{bmatrix}.
\]

The logistic response function relates the linear predictor \( \eta_t = Z \alpha_t \) to the probability vector \( \pi_t \) with elements

\[
\pi_{it} = \frac{\exp(\theta_t - \beta_i)}{1 + \exp(\theta_t - \beta_i)}.
\]  

Equation 4.9 can be seen as a dynamic variant of the Rasch model. Note that this is the form of the Rasch model without the so-called item-invariant discrimination parameter (see Hambleton & Swaminathan, 1985, p. 47). For the random walk process, the transition matrix \( F \) is simply the 5 × 5 identity matrix, \( I_5 \). For the white noise process, \( F \) is the same except that its first element \( F_{1,1} \) is zero. For the autoregressive process, the first element \( F_{1,1} \) is equal to \( \phi_1 \). The selection matrix \( R \) reduces to a vector \( r \), because we have only a single time-varying parameter \( (\theta_t) \), and is given by

\[
r = \begin{bmatrix}
1 \\
0 \\
0 \\
0 \\
0 \\
\end{bmatrix}
\]

Estimation of the latent process is discussed next.

### 4.5 Estimation

The iteratively weighted Kalman filter and smoother (KFS) as described in Fahrmeir and Wagenpfeil (1997) is used to obtain estimates of the latent process \( \alpha_t \). The KFS procedure maximizes the following log-posterior distribution
of the states \( \alpha_t, t = 1, \ldots, T \),

\[
\log p(\alpha) = \sum_{t=1}^{T} \sum_{i=1}^{n} (y_{it} \log(\pi_{it}) + (1 - y_{it}) \log(1 - \pi_{it})) \\
- \frac{1}{2}(\alpha_0 - a_0)'Q_0^{-1}(\alpha_0 - a_0) \\
- \frac{1}{2} \sum_{t=1}^{T} (\alpha_t - F_t\alpha_{t-1})'R_tQ_t^{-1}R_t'(\alpha_t - F_t\alpha_{t-1}),
\]

(4.10)

For the white noise process, the initial state contribution can be eliminated from the log-posterior. Also, in the third part of Equation 5.11, only the time-varying elements (\( \theta_t \)) of the state vector contribute to the log-posterior. However, this contribution differs for the three models that are used here, and therefore, the above representation is retained. In the presentation of the KFS procedure, the parameters or values in \( a_0, Q_0, Z, F, R, \) and \( Q \) are considered to be either known or fixed. In each iteration \( i \) of the KFS procedure, evaluation values for the state process are needed, which are denoted by \( \tilde{a}^i = (\tilde{a}^i_1, \tilde{a}^i_2, \ldots, \tilde{a}^i_T)' \). Filtering and smoothing then proceeds as follows.

### 4.5.1 Filtering

First, the filter is initialized by

\[
a_{0|0} = a_0 \quad \text{and} \quad V_{0|0} = R_0Q_0R_0'.
\]

The extended Kalman filter consists of two recursive steps, a prediction and correction step, which are taken consecutively. The prediction step is described as follows

\[
a_{t|t-1} = F_t a_{t-1|t-1}, \\
V_{t|t-1} = F_t V_{t-1|t-1} F_t' + R_tQ_tR_t'.
\]

The correction step is given by the following equations

\[
V_{t|t} = (V_{t|t-1}^{-1} + B_t)^{-1}, \\
a_{t|t} = a_{t|t-1} + V_{t|t}b_t,
\]

where \( b_t \) and \( B_t \) are the so-called working score function and expected information matrix given by

\[
B_t = Z_t'D_t\Sigma_t^{-1}D_t'Z_t, \\
b_t = Z_t'D_t\Sigma_t^{-1}(y_t - h(Z_t\tilde{a}_t)) + B_t(a_{t|t-1} - \tilde{a}_t),
\]
where \( h(.) \) is the logistic response function and, for the modelling example considered in Section 4.4.2, \( D_t \) is the symmetric Jacobian matrix of the response function given by

\[
D_t = \frac{\partial h(\eta_t)}{\partial \eta_t} = \begin{bmatrix}
\pi_{t1} (1 - \pi_{t1}) & 0 & \pi_{t2} (1 - \pi_{t2}) & 0 \\
0 & \pi_{t3} (1 - \pi_{t3}) & 0 & \pi_{t4} (1 - \pi_{t4}) \\
0 & 0 & \pi_{t3} (1 - \pi_{t3}) & 0 \\
0 & 0 & 0 & \pi_{t4} (1 - \pi_{t4})
\end{bmatrix},
\]

and \( \Sigma_t \) is the covariance matrix of \( y_t \) given by

\[
\Sigma_t = \begin{bmatrix}
\pi_{t1} (1 - \pi_{t1}) & 0 & \pi_{t2} (1 - \pi_{t2}) & 0 \\
0 & \pi_{t3} (1 - \pi_{t3}) & 0 & \pi_{t4} (1 - \pi_{t4}) \\
0 & 0 & \pi_{t3} (1 - \pi_{t3}) & 0 \\
0 & 0 & 0 & \pi_{t4} (1 - \pi_{t4})
\end{bmatrix}.
\]

Both \( D_t \) and \( \Sigma_t \) are evaluated at \( \tilde{\alpha}_i \). Note that \( D_t \) and \( \Sigma_t \) are equal for our modelling example, although, in general, this is not necessarily the case.

### 4.5.2 Smoothing

The fixed interval smoother is a backward procedure to obtain state estimates \( a_{t-1|T} \) utilizing the information of the complete time series. For \( t = T, \ldots, 2 \), we obtain

\[
a_{t-1|T} = a_{t-1|t-1} + G_t(a_{t|T} - a_{t|t-1}),
\]

\[
V_{t-1|T} = V_{t-1|t-1} + G_t(V_{t|T} - V_{t|t-1})G_t',
\]

where

\[
G_t = V_{t-1|t-1} F_t V_{t-1|t-1}'.
\]

After each application of the filter and smoother, the state evaluation values are updated with the results of the smoother, that is, \( \tilde{\alpha}_i^{i+1} = (a_{t|T}', \ldots, a_{t|T}')' \). The filter and smoother are applied repeatedly until some convergence criterion is reached. The stopping criterion of the KFS procedure used in the present investigation is \( \max |\tilde{\alpha}_i - \alpha_i^{i-1}| < 1^{-12} \).
4.6 Examples

4.6.1 Simulated data

Data were simulated using the three models described in Section 4.4.2 with the following parameter settings. Each simulated time series has length $T = 200$. The threshold parameters for all three models are the same and given by

$$\beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{bmatrix} = \begin{bmatrix} -1.5 \\ -0.5 \\ 0.5 \\ 1.5 \end{bmatrix}.$$

For identification purposes, $\mu_\theta$ is kept fixed at zero. This parameter can then be deleted from the state vector, and associated model matrices $Z$, $F$, and $R$ are of reduced dimension. The variance of the white noise process $\theta_t$ is fixed at one, so $q = 1$. The autoregressive parameter is $\phi_1 = 0.7$, and the variance of the autoregressive process $\theta_t$ is also fixed at one, which means that $q = 1 - \phi_1^2 = 0.51$. The variance of the innovations of the random walk process is $q = 0.01$. So we have four items, a single latent factor, a single person, and a series of length $T = 200$.

The results on parameter estimation for the three simulated data examples are given in Table 4.1. The estimated $\beta$'s are rescaled, so that their mean is zero, and an estimate of $\mu_\theta$ is obtained. In this example, the item parameters are recovered best with the RW process in terms of bias, but the standard errors of the item parameters are slightly larger than for WN and AR processes. The

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>WN Est.</th>
<th>SE$^2$</th>
<th>AR Est.</th>
<th>SE</th>
<th>RW Est.</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu_\theta$</td>
<td>0.00</td>
<td>0.068 (0.086)</td>
<td>0.066 (0.148)</td>
<td>0.134 (0.276)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>-1.50</td>
<td>-1.289 (0.148)</td>
<td>-1.268 (0.152)</td>
<td>-1.447 (0.160)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>-0.50</td>
<td>-0.511 (0.134)</td>
<td>-0.532 (0.139)</td>
<td>-0.500 (0.146)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_3$</td>
<td>0.50</td>
<td>0.559 (0.136)</td>
<td>0.461 (0.140)</td>
<td>0.504 (0.153)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_4$</td>
<td>1.50</td>
<td>1.241 (0.151)</td>
<td>1.339 (0.158)</td>
<td>1.442 (0.182)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1 Estimate
2 Standard error

Table 4.1: Results on parameter estimation for simulated data
differences in standard error of $\mu_\theta$ for the three processes are remarkably large compared to those for the item parameters.

The results of filtered and smoothed probabilities, and filtered and smoothed states are presented in figures. Figure 4.1 displays the smoothed probabilities of the white noise process and Figure 4.2 displays the true and estimated latent process. Since the process is sequentially independent, the smoothed process can take on only five different values (the number of possible sumscores on the four items). This is seen in Figure 4.1, where each item has five probabilities and the probability values are dependent on the $\beta$'s. Because the estimated latent process can take on only five different values, the true process is not well tracked, which can be seen in Figure 4.2.

Figure 4.3 displays the true and estimated probabilities for the AR example.
Since the process is dependent on its first lag, the estimated latent process can take on many more values than the number of sumscores. The estimated probabilities track the true probability with reasonable accuracy. Figure 4.4 shows the true and estimated process paths. The true process is roughly tracked, although sometimes peaks or jumps seem difficult to recover.

Figures 4.5 and 4.6 show the results obtained with the random walk. Since a relatively small variance was selected for the innovations, the process is smoother than the first two examples. The estimated probabilities follow the true probabilities, although they tend to be too smooth. The divergence of the estimated latents process, however, is more severe, which is clearly seen in the middle of the time series in Figure 4.6.

4.6.2 Real data

Real data were analysed with the described techniques. We selected a single subject and a single subscale (Neuroticism) containing six items of a data set consisting of personality questionnaires containing 30 items scored on seven-point scales, administered to 22 psychology students on 90 consecutive days (Borkenau
The questionnaires were constructed as to measure the Big Five personality factors (McCrae & John, 1992). The data were dichotomized for illustrative purposes only in order to apply the dynamic logistic model.

The WN, AR, and RW processes were fitted to the observed time series. Since this example is for illustrative purposes, the parameters $q$ and $\phi_1$ were fixed at the same values as in the simulated data examples. The results on the estimation of the other parameters is displayed in Table 4.2. The estimates of $\mu_\theta$ display the same pattern as in the simulated data. The item parameter estimates show a different pattern in that with the WN process, the estimates are somewhat more spread out than with the AR process and the RW process. The standard errors

---

7Data were kindly made available by professor Borkenau.
of the item parameters do not differ much between the three processes.

Figure 4.7 shows the estimated latent WN, AR, and RW processes. On the left side, the results of the filter are displayed, and on the right side, the results of the smoother. Some observations can be made. For the WN process, there are as many values for the process as there are sumscores, but only after the smoother has been applied. The effect of smoothing is much larger for the RW process than for the other two processes in this example. The smoothed RW process displays an interesting swelling pattern which is difficult to unveil with the WN process. In the AR process, the pattern of the RW process is becoming visible.

4.7 Discussion

In the present chapter we took a closer look at the rationale for the emphasis in psychometrics on the analysis of IEV. It was found that this rationale is weak and that arguments for analysis of IAV are too easily brushed aside. We provided arguments for the development of models based on IAV. The question of the existence of any lawful relationship between analysis of IEV and IAV was addressed
and it was argued that there are criteria for the existence of such a relationship. These criteria, however, are very strict and are met only when the processes concerned are ergodic. Since, in practice, little is known about the relation between analysis of IEV and IAV, and thus about ergodicity of the processes concerned in psychometrics, investigation of this relation is important. First, however, reliable methods have to be developed for analysis of IAV. The present chapter attempted to provide an outline of methods for analysing single-subject dichotomous time series.

An advantage of the presented modelling approach is that models for polychotomous responses can be easily obtained after appropriate adjustments. In addition, it can be investigated if several persons can be analysed with the same model with equal parameter settings, i.e., if measurement invariance holds. How-
ever, the results of the simulated and real data examples indicate that the discussed modelling outline requires further investigation. Important topics in this investigation are the development of estimation methods for variances and autoregressive parameters. Fahrmeir and Wagenpfeil (1997) discuss a procedure to estimate $a_0$, $Q_0$, and $Q$, but little is known about its behavior. Finally, in order to perform a full analysis of real data, methods to evaluate the fit of the discussed types of models become a necessary tool. Such methods await development and investigation.

<table>
<thead>
<tr>
<th>Item</th>
<th>Parameter</th>
<th>WN</th>
<th>SE</th>
<th>AR</th>
<th>SE</th>
<th>RW</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\mu_0$</td>
<td>0.276</td>
<td>(0.123)</td>
<td>0.319</td>
<td>(0.222)</td>
<td>0.637</td>
<td>(0.319)</td>
</tr>
<tr>
<td>Irritable (+)</td>
<td>$\beta_1$</td>
<td>-0.117</td>
<td>(0.214)</td>
<td>-0.109</td>
<td>(0.207)</td>
<td>-0.100</td>
<td>(0.202)</td>
</tr>
<tr>
<td>Emotionally stable (-) $^1$</td>
<td>$\beta_2$</td>
<td>0.211</td>
<td>(0.217)</td>
<td>0.193</td>
<td>(0.210)</td>
<td>0.176</td>
<td>(0.205)</td>
</tr>
<tr>
<td>Calm (-)</td>
<td>$\beta_3$</td>
<td>0.046</td>
<td>(0.215)</td>
<td>0.041</td>
<td>(0.208)</td>
<td>0.037</td>
<td>(0.203)</td>
</tr>
<tr>
<td>Bad-tempered (+)</td>
<td>$\beta_4$</td>
<td>0.799</td>
<td>(0.230)</td>
<td>0.738</td>
<td>(0.224)</td>
<td>0.679</td>
<td>(0.219)</td>
</tr>
<tr>
<td>Resistant (-)</td>
<td>$\beta_5$</td>
<td>-0.714</td>
<td>(0.216)</td>
<td>-0.655</td>
<td>(0.210)</td>
<td>-0.601</td>
<td>(0.204)</td>
</tr>
<tr>
<td>Vulnerable (+)</td>
<td>$\beta_6$</td>
<td>-0.225</td>
<td>(0.214)</td>
<td>-0.208</td>
<td>(0.207)</td>
<td>-0.191</td>
<td>(0.202)</td>
</tr>
</tbody>
</table>

$^1$ Items with a minus are negatively formulated and therefore recoded.
Figure 4.7: Filter (left) and smoother (right) results of fitting white noise (top), autoregressive (middle), and random walk (bottom) processes to single subject neuroticism data.