An Alternating Poisson Process model for estimating prevalence and incidence from

interval recording data

James E. Pustejovsky and Daniel M. Swan

The University of Texas at Austin

Author Note

James E. Pustejovsky, Department of Educational Psychology, University of Texas
at Austin. Daniel M. Swan, Department of Educational Psychology, University of Texas
at Austin.

Address correspondence to James E. Pustejovsky, Department of Educational
Psychology, University of Texas at Austin, 1 University Station D5800, Austin, TX
78712. Email: pusto@austin.utexas.edu.

Abstract

Keywords: behavioral observation; interval recording; alternating Poisson process; Markov chain

An Alternating Poisson Process model for estimating prevalence and incidence from interval recording data

Measurements derived from systematic, direct observation of human behavior are used in many areas of psychological and educational research. For example, direct observation of student classroom behavior is a primary component of several existing instruments for screening and diagnosis of emotional and behavioral problems (**?**); measures based on direct observation of childrens' challenging behavior in home settings have been employed as pre- and post-test measures in randomized trials of behavioral interventions (e.g., **?**); and direct observation of infant-parent interaction patterns is employed in studies of child development (Mann, Ten Have, Plunkett, & Meisels, 1991) and cross-cultural differences (**?**). Direct observation also plays a particularly prominent role in single-case research, where it is used to measure changes in behavior over time and assess individual responses to intervention (**?**).

Systematic direct observation procedures require that the behavior of interest have a clear operational definition, so that its occurrence or absence can be judged at a given point in time. In forming such an operationally definition, is useful to distinguish between behaviors that are *events*, where each occurrence is of negligible duration, versus behaviors that are *states*, where individual episodes of behavior have positive duration (J. Altmann, 1974). The primary characteristic of an event behavior is its incidence, or frequency of occurrence per unit time. The two primary characteristics of a state behavior are its incidence–the frequency with which new episodes of the behavior begin, per unit time–and its prevalence, or the proportion of time that it occurs.

Explain observation session

Several commonly used systems for collecting behavioral data do not capture a complete record of the behavior during an observation session, but rather involve making observations only intermittently. Among intermittent recording procedures, the three main procedures are momentary time sampling, partial interval recording, and

whole interval recording. In all three methods, an observation session is divided into a fixed number of equally spaced intervals, of perhaps 10 or 15 s in length, and a binary data-point is recorded for each interval. The systems differ only in the rule for scoring each interval. Using momentary time sampling (MTS), an interval is scored as a one if a behavioral event is happening during the final moment of the interval (and is otherwise scored as a zero). Using partial interval recording (PIR), an interval is scored as a one if the behavior occurs at any point during the interval. Using whole interval recording (WIR), an interval is scored as a one only if the behavior occurs for the entire duration of the interval. In some PIR and WIR systems, a small length of time is left between each interval so that the observer does not have to maintain continuous attention.

In many applications, the interval-by-interval data generated by these recording systems is summarized by the proportion of intervals scored as a one. Often, only this summary proportion is used for later analysis, where it is interpretted as a measure of prevalence. Under what circumstances is it reasonable to reduce the data to the summary proportion? We argue that this depends on which recording system is used to collect the measurements.

Under quite general modeling assumptions, the proportion of MTS intervals is an unbiased estimate of prevalence (Rogosa & Ghandour, 1991). Thus, reducing the data to the summary proportion may be quite reasonable if the investigator's interest is solely in the prevalence of the behavior. However, simple summaries of MTS data do not provide a clear measure of the behavior's incidence. Even if not of substantive interest, estimates of incidence are are necessary for assessing the magnitude of measurement error in the prevalence estimates.

Brown, Solomon, and Stephens (1977) described a method for estimating both prevalence and incidence from MTS data. Their approach was to first posit a stochastic process for the underlying stream of behavior, as perceived by the observer, and then to consider the implications of observing the process intermittently via an MTS system.

The particular model they considered was an Alternating Poisson Process, which is a simple, two-state continuous time Markov chain where transitions between states follow exponential distributions. The Alternating Poisson Process implies that interval-by-interval MTS scores follow a discrete-time Markov chain. Under the model, Brown et al. (1977, see also Griffin & Adams, 1983) provided closed-form expressions for the maximum likelihood estimators of prevalence and incidence and the associated asymptotic covariance matrix.

Unlike MTS, PIR and WIR systems do not produce clearly interpretable summary measurements. Rather, the PIR summary proportion systematically over-estimates prevalence and the WIR summary proportion systematically under-estimates prevalence; in both cases, the extent of the bias depends on the incidence of the behavior as well as operational features of the recording system (Kraemer, 1979; Rogosa & Ghandour, 1991), making the construct interpretation of such data quite difficult. Consequently, methodologists have long argued against the use of PIR and WIR systems (J. Altmann, 1974; Lane & Ledford, 2014; Mann et al., 1991, cf.). Despite such objections, the systems remain in common use, particularly as part of behavioral time series designs and single-case research (Lane & Ledford, 2014; Mudford, Taylor, & Martin, 2009; Rapp et al., 2007).

Cite example about incorrect inferences based on PIR?

Little previous research has considered methods of analyzing PIR and WIR data beyond using the summary proportion. For PIR data, S. A. Altmann and Wagner (1970) proposed a transformation of the summary proportion as an estimate of incidence, motivated by a model in which the new behavioral episodes follow a Poisson process. However, their model applies only to "event" behaviors (J. Altmann, 1974). Suen and Ary (1986, 1989) have proposed a method for obtaining estimates of prevalence and incidence from PIR data, provided that the behavior stream conforms to certain conditions. However, their proposed procedure is not motivated by any explicit data-generating process, and later simulation studies reported that the method produces

badly biased estimates (Rogosa & Ghandour, 1991, sec. 5.2).

Add short discussion of Pustejovsky and Swan (2014).

This paper examines models for PIR and WIR data, from which principled estiamtes of prevalence and incidence can be obtained. Following Brown et al. (1977), we use an Alternating Poisson Process for the underlying behavior stream to derive a model for the interval-by-interval scores. Under this model, maximum likelihood estimators for prevalence and incidence can be obtained using conventional numerical techniques (although they do not have closed-form expression). To remedy some problems with the maximum likelihood estimators, we introduce penalized likelihood estimators that have better operating characteristics and that can be tailored to express prior information about the behavioral parameters. We use a parametric bootstrap procedure for obtaining interval estimates.

Brief discussion of penalized likelihood estimation in other contexts.

The remainder of the paper is organized as follows. In the next section, we introduce the Alternating Poisson Process model for the behavior stream and lay out common notation. In the following section, we review the estimators described by Brown et al. (1977) for MTS data and propose alternative estimators based on penalized likelihood. In the following section, we develop the Alternating Poisson Process model to handle PIR data, then study penalized likelihood estimators for prevalence and incidence. Finally, we present a brief empirical application and discuss limitations and future research directions.

Revise this.

Alternating Poisson Process models

The Alternating Poisson Process is a stochastic model that can be used to describe a stream of behavior, as it is perceived in time. The model applies to state behaviors, where each episode has a non-negligible duration. It is assumed that the behavior is either occurring or not occurring at any given point in time, so that the perceived behavior stream can be described in terms of two components: sequentially ordered,

non-overlapping episodes of behavior, which we will call *event durations*, and spans of time in between episodes, which we will call *interim times*. Let $\{Z(t), 0 \leq t\}$ denote the state of the behavior stream over the course of an observation session, where $Z(t) = 1$ indicates that an event is occurring at time t and $Z(t) = 0$ otherwise.

In the Alternating Poisson Process, it is assumed that event durations and interim times are mutually independent, random quantities, that the event durations follow an exponential distribution with mean $\mu > 0$, and that the interim times follow an exponential distribution with mean $\lambda > 0$. Under the model, the prevalence of the behavior is equal to the ratio of μ to the sum of μ and λ and the incidence of the behavior is equal to the reciprocal of the sum of μ and λ. We will denote prevalence by ϕ, where $0 < \phi < 1$, and incidence by ζ, where $\zeta > 0$.

Finally, we assume that the process is in equilibrium, with $\Pr(Y(0) = 1) = \phi$. This assumption implies that there is a constant marginal probability of observing an event at any given point in time.

> Expand further– when would this not apply/be inappropriate?

The Alternating Poisson Process is a special case of a continuous time Markov chain, and thus has the Markov property that the the future state of the behavior depends only on the current state, but not on the past history of the behavior. More precisely, the probability that a behavior will be occurring t seconds into the future is independent of the state of the behavior for $0 \leq r < s$:

$$\Pr\left[Z(s+t) = 1 \,\middle|\, Z(s) = a, Z(r) : 0 \leq r < s\right] = \Pr\left[Z(s+t) = 1 \,\middle|\, Z(s) = a\right] \quad (1)$$

for $a = 0, 1$ and $s, t \geq 0$ (Kulkarni, 2010, Thm. 6.1). The assumption that the process is in equilibrium further implies that the probability that a behavior will be occurring t seconds into the future does not depend on the current time, i.e.,

$$\Pr\left[Z(s+t) = 1 \,\middle|\, Z(s) = a\right] = \Pr\left[Z(t) = 1 \,\middle|\, Z(0) = a\right]. \quad (2)$$

Let $p_a(t)$ denote the conditional probability that an event will be occurring t seconds into the future, given that the behavior is currently in state a, for $a = 0, 1$. These

conditional probabilities can be expressed as follows:

$$p_0(t) = \Pr(Z(t) = 1 | Z(0) = 0) = \phi \left[1 - \exp\left(\frac{-t\zeta}{\phi(1-\phi)} \right) \right]$$

$$p_1(t) = \Pr(Z(t) = 1 | Z(0) = 1) = \phi + (1-\phi) \exp\left(\frac{-t\zeta}{\phi(1-\phi)} \right)$$

(3)

(Kulkarni, 2010, Eq. 6.17).

Momentary Time Sampling

Consider observing a behavior stream generated by the Alternating Poisson Process and recording observations using momentary time sampling with $K + 1$ recording times, equally spaced at intervals of length c. Denote the recorded data by the sequence of binary indicator variables $X_0, X_1, ..., X_K$. The MTS interval data are a record of the state of the behavior stream process at fixed moments in time: $X_k = Z(ck)$ for $k = 0, ..., K$.

Brown et al. (1977) demonstrated that MTS data follow a two-state, discrete-time Markov chain process with transition probabilities $Pr(X_k = 1 | X_{k-1} = a) = p_a(c)$ and $Pr(X_k = 0 | X_{k-1} = a) = 1 - p_a(c)$ for $a = 0, 1$. Therefore, sufficient statistics for the process are given by the table counting the number of transitions with $(X_{k-1} = a, X_k = b)$ for $a, b = 0, 1$ and $k = 1, ..., K$; let $n_{ab} = \sum_{k=1}^{K} I(X_{k-1} = a, X_k = b)$. Conditioning on X_0, the log-likelihood of MTS data is then given by

$$l_{MTS}(\phi, \zeta) = n_{01} \log \phi + n_{10} \log (1 - \phi)$$

$$+ (n_{01} + n_{10}) \log \left[1 - \exp\left(\frac{-\zeta c}{\phi(1-\phi)} \right) \right]$$

$$+ n_{00} \log \left[1 - \phi + \phi \exp\left(\frac{-\zeta c}{\phi(1-\phi)} \right) \right]$$

$$+ n_{11} \log \left[\phi + (1 - \phi) \exp\left(\frac{-\zeta c}{\phi(1-\phi)} \right) \right].$$

(4)

Let $\hat{p}_0 = n_{01} / (n_{00} + n_{01})$ and $\hat{p}_1 = n_{11} / (n_{10} + n_{11})$. Brown et al. (1977) showed that the maximum likelihood estimate of ζ exists only when $\hat{p}_0 < \hat{p}_1$. When this condition holds, the maximum likelihood estimates of ϕ and ζ are given by

$$\hat{\phi}_{MTS} = \frac{\hat{p}_0}{\hat{p}_0 + 1 - \hat{p}_1} \quad \text{and} \quad \hat{\zeta}_{MTS} = \frac{-\hat{p}_0 (1 - \hat{p}_1) \log(\hat{p}_1 - \hat{p}_0)}{c (\hat{p}_0 + 1 - \hat{p}_1)^2}.$$

(5)

If $\hat{p}_0 \geq \hat{p}_1$, then the maximum likelihood estimator of ϕ is still well-defined and equal to $\hat{\phi}_{MTS} = (n_{01} + n_{11})/K$.

The probability that the maximum likelihood estimators are undefined or fall outside of the parameter space is not trivial, even when K is relatively large. Note that in order for the estimates to fall strictly within the parameter space, both a 0-1 transition and a 1-0 transition must be observed, so that $\hat{p}_0 > 0$, $\hat{p}_1 < 1$. Table 1 reports the proportion of 2000 simulated samples in which $0 < \hat{\phi}_{MTS} < 1$ and $0 < \hat{\zeta}_{MTS} < \infty$, with $K = 40$ and ζ scaled in terms of the interval length. Values of $\phi > 0.5$ are omitted because the behavior of the MTS estimators is symmetric about $\phi = 0.5$. The proportion of estimates falling within in the parameter space decreases as prevalence becomes more extreme and as incidence becomes very infrequent or very frequent. The high proportion of samples in which the estimators are undefined is a substantial drawback to the use of maximum likelihood.

Partial Interval Recording

Consider observing a behavior stream generated by the Alternating Poisson Process and recording observations using partial interval recording. Suppose that one observes K intervals, where each interval includes c seconds of active observation time

Table 1

Proportion of 2000 simulated MTS samples ($K = 40$) in which $0 < \hat{\phi}_{MTS} < 1$ and $0 < \hat{\zeta}_{MTS} < \infty$.

	$\zeta = 0.02$	0.05	0.1	0.2	0.25	0.4	0.5
$\phi = 0.1$		0.43	0.63	0.64	0.45	0.37	0.29	0.26
0.2		0.45	0.80	0.90	0.84	0.79	0.59	0.49
0.3		0.46	0.84	0.96	0.96	0.94	0.73	0.64
0.4		0.43	0.87	0.98	0.99	0.97	0.82	0.71
0.5		0.45	0.88	0.99	1.00	0.98	0.84	0.71

followed by d seconds of recording time. Let time $t_k = (k-1)(c+d)$ denote the beginning of interval k. Let U_k indicate the PIR score from interval k, corresponding to the time from t_k to $t_k + c$. Following the PIR system, $U_k = 1$ if is the behavior occurs at any point during the active portion of interval, and $U_k = 0$ otherwise. In terms of the behavior stream process,

$$U_k = I\left[0 < \int_0^c Z\left(t_k + s\right) ds\right] \qquad (6)$$

for $k = 1, ..., K$, where $\int_0^c$ denote the definite integral over the half-open interval $[0, c)$.

Under the assumptions of the Alternating Poisson Process, the joint distribution of $U_1, ..., U_K$ can be derived as follows. Let $\psi_k, k = 2, ..., K$ denote the probability than the behavior is occurring at time $t_k = (k-1)(c+d)$, given the partial interval record up to that time. For ease of notation, set $\psi_1 = \phi$, which follows from the assumption that the process is in equilibrium. We show in Appendix A that

$$\psi_k = \Pr\left[Z(t_k) = 1 \mid U_1, ..., U_{k-1}\right]$$
$$= \left[\frac{\psi_{k-1} p_1(c+d) + (1 - \psi_{k-1})\left[p_0(c+d) - p_0(d) \exp\left(\frac{-\zeta c}{1-\phi}\right)\right]}{1 - (1 - \psi_{k-1}) \exp\left(\frac{-\zeta c}{1-\phi}\right)}\right]^{u_{k-1}} [p_0(d)]^{(1-u_{k-1})}.$$
$$(7)$$

Note that $Z(t_k) = 1$ implies that $U_k = 1$ with certainty, while

$$\Pr\left(U_k = 1 \mid Z(t_k) = 0\right) = 1 - \exp\left(\frac{-\zeta c}{1-\phi}\right).$$

It follows from the Markov property of the Alternating Poisson Process that

$$\Pr\left(U_k = 1 \mid U_1, ..., U_{k-1}\right) = \psi_k \Pr\left(U_k = 1 \mid Y(t_k) = 1\right)) + (1 - \psi_k) \Pr\left(U_k = 1 \mid Y(t_k) = 0\right))$$
$$= 1 - (1 - \psi_k) \exp\left(\frac{-\zeta c}{1-\phi}\right).$$

The joint distribution of $U_1, ..., U_K$ can therefore be expressed as

$$\Pr\left(U_1 = u_1, ..., U_K = u_K\right) = \Pr\left(U_1 = u_1\right) \prod_{k=2}^K \Pr\left(U_k = u_k \mid U_1, ..., U_{k-1}\right)$$
$$= \prod_{k=1}^K \left[1 - (1 - \psi_k) \exp\left(\frac{-\zeta c}{1-\phi}\right)\right]^{u_k} \left[(1 - \psi_k) \exp\left(\frac{-\zeta c}{1-\phi}\right)\right]^{(1-u_k)}.$$

The log-likelihood of ϕ and ζ, given observed PIR data $u_1, ..., u_K$, is

$$l_{PIR}(\phi, \zeta) = \sum_{k=1}^{K} u_k \ln \left[1 - (1 - \psi_k) \exp \left(\frac{-\zeta c}{1 - \phi} \right) \right] + (1 - u_k) \left[\ln(1 - \psi_k) - \frac{\zeta c}{1 - \phi} \right]. \quad (8)$$

Maximum likelihood estimates are obtained by maximizing l_{PIR} using a numerical algorithm. Because the conditional probabilities $\psi_1, ..., \psi_K$ are defined recursively, it is cumbersome and computationally expensive to evaluate the score function corresponding to this likelihood. The simulation study reported in the next subsection therefore uses the Nelder-Mead algorithm, which does not require evaluation of the score function, with the likelihood expressed in terms of the unbounded parameters $\text{logit}(\phi)$ and $\log(\zeta)$.

Just as with MTS data, maximum likelihood estimates based on PIR data do not always fall within the parameter space. Table 2 reports the proportion of 2000 simulated samples in which the MLEs based on PIR data are within the parameter space, again with $K = 40$ and ζ scaled in terms of the interval length. Because we use numerical maximization, the results of the maximization routine are never precisely on the borders

Table 2

Proportion of 2000 simulated PIR samples ($K = 40$) in which $|logit(\hat{\phi})| < 8$ and $|log(\hat{\zeta})| < 8$.

	$\zeta = 0.02$	0.05	0.1	0.2	0.25	0.4	0.5
$\phi = 0.1$	0.59	0.83	0.94	0.94	0.90	0.78	0.72
0.2	0.67	0.91	0.98	0.99	0.98	0.92	0.84
0.3	0.74	0.94	0.99	1.00	1.00	0.94	0.87
0.4	0.78	0.96	1.00	1.00	0.99	0.93	0.86
0.5	0.77	0.96	0.99	0.99	0.98	0.89	0.80
0.6	0.75	0.94	0.97	0.96	0.92	0.79	0.67
0.7	0.68	0.87	0.92	0.85	0.78	0.56	0.38
0.8	0.60	0.76	0.77	0.56	0.42	0.16	0.09
0.9	0.47	0.48	0.32	0.08	0.04	0.00	0.00

of the parameter space. We therefore use boundaries of $|\text{logit } \hat{\phi}| < 8$ and $|\log \hat{\zeta}| < 8$ to define the edges of the parameter space.

Discuss table.

In addition to returing estimates that are on the edges of the parameter space, the maximum likelihood estimators have the further disadvantage of being somewhat sensitive to initialization values. The likelihood surface becomes very flat when the PIR scores are near ceiling or floor levels, making it difficult to numerically identify the maximum. For the implementation of the estimators in the accompanying R package, the consequence is that drastically different estimates can be returned depending on the initialization values of the algorithm. Together with the possibility of obtaining estimates on the edges of the parameter space, the algorithmic instability of the maximum likelihood estimators motivates our investigation of alternative estimators that incorporate penalty functions.

Whole Interval Recording

Consider observing a behavior stream generated by the Alternating Poisson Process and recording observations using whole interval recording. As with PIR, suppose that one observes K intervals, where each interval includes c seconds of active observation time followed by d seconds of recording time. Let W_k indicate the WIR score from interval k, corresponding to the time from t_k to $t_k + c$. Following the WIR system, $W_k = 1$ if is the behavior occurs for the duration of the active portion of interval, and $W_k = 0$ otherwise. Formally,

$$W_k = I\left[c = \int_0^c Z\left(t_k + s\right) ds\right] \tag{9}$$

for $k = 1, ..., K$.

Using the WIR system to score a state behavior is logically equivalent to using PIR to score the absence of the behavior. WIR data can therefore be modeled just as PIR data, after an appropriate change of parameters. Specifically, the log-likelihood for WIR data under the Alternating Poisson Process can be written in terms of the

log-likelihood for PIR data as

$$l_{WIR}(\phi, \zeta | W_1 = w_1, ..., W_k = w_k) = l_{PIR}(1 - \phi, \zeta | U_1 = 1 - w_1, ..., U_K = 1 - w_k). \quad (10)$$

The equivalence of the two system implies that maximum likelihood and penalized likelihood estimators based on WIR data can be obtained using the algorithms developed for PIR.

Penalized likelihood estimators

Potential priors:

- Independent $\Gamma(k, \theta)$ priors on μ, λ scale, which translates to $\phi \sim Beta(k_\mu, k_\lambda)$ and $1/\zeta \sim \Gamma(k_\mu + k_\theta, \theta)$. Need to have $k_\mu, k_\lambda > 1$ in order for estimates to remain within parameter space.

- Independent normal priors on $\text{logit}\phi, \log \zeta$.

Finite-sample performance: MTS

Finite-sample performance: PIR

Application

Discussion

References

Altmann, J. (1974). Observational study of behavior: Sampling methods. *Behaviour*, *49*(3/4), 227–267.

Altmann, S. A., & Wagner, S. S. (1970). Estimating rates of behavior from Hansen frequencies. *Primates*, *11*(2), 181–183. doi: 10.1007/BF01731143

Brown, M., Solomon, H., & Stephens, M. A. (1977). Estimation of Parameters of Zero-One Processes by Interval Sampling. *Operations Research*, *25*(3), 493–505.

Griffin, B., & Adams, R. (1983). A parametric model for estimating prevalence, incidence, and mean bout duration from point sampling. *American Journal of Primatology*, *4*(3), 261–271. doi: 10.1002/ajp.1350040305

Kraemer, H. C. (1979). One-zero sampling in the study of primate behavior. *Primates*, *20*(2), 237–244.

Kulkarni, V. G. (2010). *Modeling and Analysis of Stochastic Systems.* Boca Raton, FL: Chapman & Hall/CRC.

Lane, J. D., & Ledford, J. R. (2014). Using interval-based systems to measure behavior in early childhood special education and early intervention. *Topics in Early Childhood Special Education.* doi: 10.1177/0271121414524063

Mann, J., Ten Have, T. R., Plunkett, J. W., & Meisels, S. J. (1991). Time sampling: A methodological critique. *Child Development*, *62*(2), 227–241.

Mudford, O. C., Taylor, S. A., & Martin, N. T. (2009). Continuous recording and interobserver agreement algorithms reported in the Journal of Applied Behavior Analysis (1995-2005). *Journal of Applied Behavior Analysis*, *42*(1), 165–169. doi: 10.1901/jaba.2009.42-165

Pustejovsky, J. E., & Swan, D. M. (2014). *Four methods for analyzing partial interval recording data, with application to single-case research.* Paper presented at the annual meeting of the American Educational Research Association, Philadelphia, PA.

Rapp, J. T., Colby-Dirksen, A. M., Vollmer, T. R., Roane, H. S., Lomas, J., Britton, L. N., & Colby, A. M. (2007). Interval recording for duration events: A re-evaluation. *Behavioral Interventions*, *22*, 319–345.

Rogosa, D., & Ghandour, G. (1991). Statistical models for behavioral observations. *Journal of Educational Statistics*, *16*(3), 157–252.

Suen, H. K., & Ary, D. (1986, March). A post hoc correction procedure for systematic errors in time-sampling duration estimates. *Journal of Psychopathology and Behavioral Assessment*, *8*(1), 31–38. doi: 10.1007/BF00960870

Suen, H. K., & Ary, D. (1989). *Analyzing Quantitative Behavioral Observation Data.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Appendix

Derivation of PIR model

The joint distribution of PIR observations depends on the conditional probabilities $\psi_k = \Pr\left[Z(t_k) = 1 \,|U_1, ..., U_{k-1}\right]$. This appendix provides a derivation of Expression (7) in terms of the parameters of the Alternating Poisson Process. Note that $U_{k-1} = 0$ implies that $Z(t_k + c) = 0$. It follows from the Markov property that

$$\Pr\left(Z(t_k) = 1 \,|U_1 = u_1, ..., U_{k-2} = u_{k-2}, U_{k-1} = 0\right)$$
$$= \Pr\left(Z(t_k) = 1 \,|Z(t_{k-1} + c) = 0\right) = p_0(d).$$

Next, observe that

$$\Pr\left(Z(t_k) = 1, U_{k-1} = 1 \,|Z(t_{k-1}) = 1\right) = \Pr\left(Z(t_k) = 1 \,|Z(t_{k-1}) = 1\right) = p_1(c + d)$$

and

$$\Pr\left(Z(t_k) = 1, U_{k-1} = 1 \,|Z(t_{k-1}) = 0\right)$$
$$= \int_0^c p_1(c - t) \frac{\zeta \exp(-t\zeta/(1 - \phi))}{(1 - \phi)} dt$$
$$= \phi \left[1 - \exp\left(\frac{-\zeta(c + d)}{\phi(1 - \phi)}\right) - \exp\left(\frac{-\zeta c}{1 - \phi}\right) + \exp\left(\frac{-\zeta(\phi c + d)}{\phi(1 - \phi)}\right)\right]$$
$$= p_0(c + d) - p_0(d) \exp\left(\frac{-\zeta c}{1 - \phi}\right).$$

It follows further that

$$\Pr\left(Z(t_k) = 1, U_{k-1} = 1 \,|U_1, ..., U_{k-2}\right)$$
$$= \psi_{k-1} p_1(c + d) + (1 - \psi_{k-1}) \left[p_0(c + d) - p_0(d) \exp\left(\frac{-\zeta c}{1 - \phi}\right)\right]$$

and

$$\Pr\left(Z(t_k) = 1 \,|U_1 = u_1, ..., U_{k-2} = u_{k-2}, U_{k-1} = 1\right)$$
$$= \frac{\psi_{k-1} p_1(c + d) + (1 - \psi_{k-1}) \left[p_0(c + d) - p_0(d) \exp\left(\frac{-\zeta c}{1 - \phi}\right)\right]}{1 - (1 - \psi_{k-1}) \exp\left(\frac{-\zeta c}{1 - \phi}\right)}.$$

Thus, ψ_k can be written as a function of ψ_{k-1} and u_{k-1}, as given in (7).