Observation procedures and Markov Chain models for estimating the prevalence and incidence of a state behavior

James E. Pustejovsky and Daniel M. Swan
The University of Texas at Austin

Austin H. Johnson
University of Connecticut

April 16, 2015

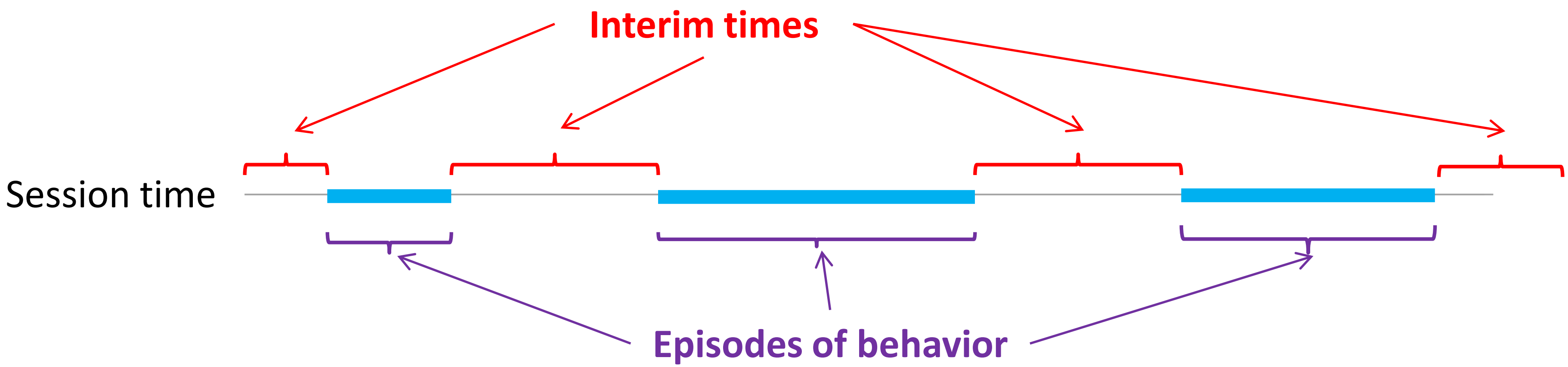
Slides at http://bit.ly/1NQIFEY

Direct observation of a state behavior

- Applications in education research
  - Assessment of student attention/disruptive behavior
  - Early childhood development
  - Single-case research for evaluating interventions for individuals with developmental disabilities

- State behavior: a behavior where individual episodes have positive duration.
  - Prevalence: proportion of time that the behavior occurs
  - Incidence: rate at which new episodes of behavior begin

The behavior stream



How to record data during direct observation of a behavior?

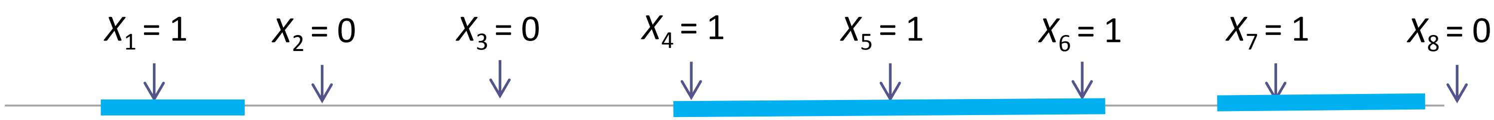
Observation recording procedures

- Continuous recording
  - Produce rich data, amenable to sophisticated modeling
  - Effort-intensive

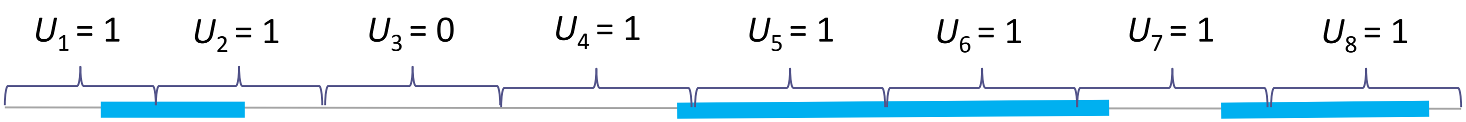
- Discontinuous recording procedures
  - Momentary Time Sampling (MTS)
  - Partial Interval Recording (PIR)
  - Whole Interval Recording (WIR)

Discontinuous recording procedures

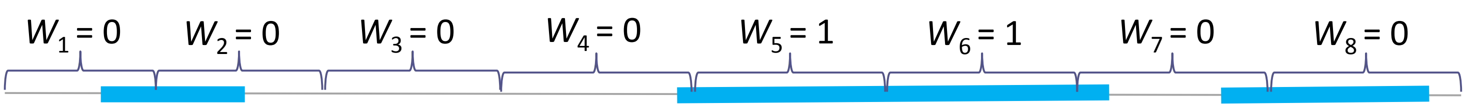
- Divide an observation session into many short intervals of equal length.
- Record a **binary score** for each interval.
- MTS: score 1 if behavior is occurring at the **end of the interval**.



- PIR: score 1 if behavior occurs **at any point** during the interval.



- WIR: score 1 if behavior occurs **for the entire interval**.



Estimation: Momentary time sampling

- MTS summary proportion is unbiased estimate of prevalence.
  - But how to estimate incidence?
  - But how to estimate standard error of measurement?

- Brown, Solomon, & Stephens (1977)
  - Show that if the behavior stream follows an Alternating Poisson Process, then MTS data follow a simple, two-state Markov Chain.
  - Provide maximum likelihood estimators for prevalence and incidence.

Estimation: Partial/whole interval recording

- PIR and WIR summary proportions are **biased** estimates of prevalence.
  - Bias depends on incidence and interval length (Kraemer, 1979; Pustejovsky, 2014).

- Alternating Poisson Process model
  - We derive expressions for the likelihood of PIR and WIR data assuming that the behavior stream follows an Alternating Poisson Process.
  - Maximum likelihood estimators, penalized maximum likelihood estimators for prevalence and incidence (using numerical optimization).
    - Prevalence PLE is much more accurate than summary proportion.
    - Relatively long sessions are needed to obtain low-bias estimates of prevalence and incidence.

A novel recording procedure: Augmented interval recording (AIR)

- Combine MTS, PIR, and WIR to obtain more efficient estimates.
- Fewer, longer intervals, with **two** binary scores per interval.

- Under the Alternating Poisson Process, AIR data follow an 4-state discrete-time Markov Chain.
  - Maximum likelihood, penalized maximum likelihood estimators for prevalence and incidence (using numerical optimization)

- For a fixed session length (and double-length intervals), AIR is…
  - More efficient than PIR/WIR (both prevalence and incidence)
  - Slightly less efficient than MTS prevalence estimates
  - Much more efficient than MTS incidence estimates

Future work

- Further evaluation
  - Robustness to violations of Alternating Poisson Process assumptions
  - Feasibility/ease of using AIR in field settings
- Extending the measurement models
  - Regressions for multiple observation sessions
  - Random effects models to describe variation in prevalence and incidence across individuals
- Psychometrics (cf. Rogosa & Ghandour, 1991)
  - Use the models to develop guidance regarding use of the recording procedures, interval lengths, and session lengths.
  - Understanding structure of recording errors under each procedure

Thank you

James E. Pustejovsky

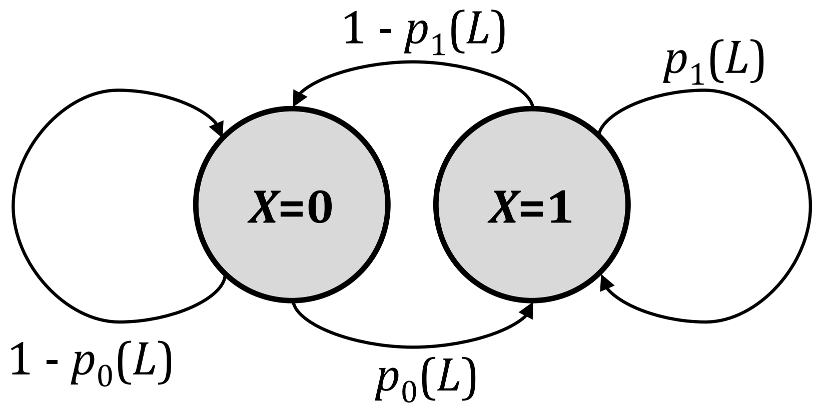
The University of Texas at Austin

Educational Psychology Department

pusto@austin.utexas.edu

Model for MTS data

- Under the alternating Poisson process, $X_1,\ldots,X_K$ follow a discrete-time Markov chain (DTMC) with two states (see e.g., Kulkarni, 2010).



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

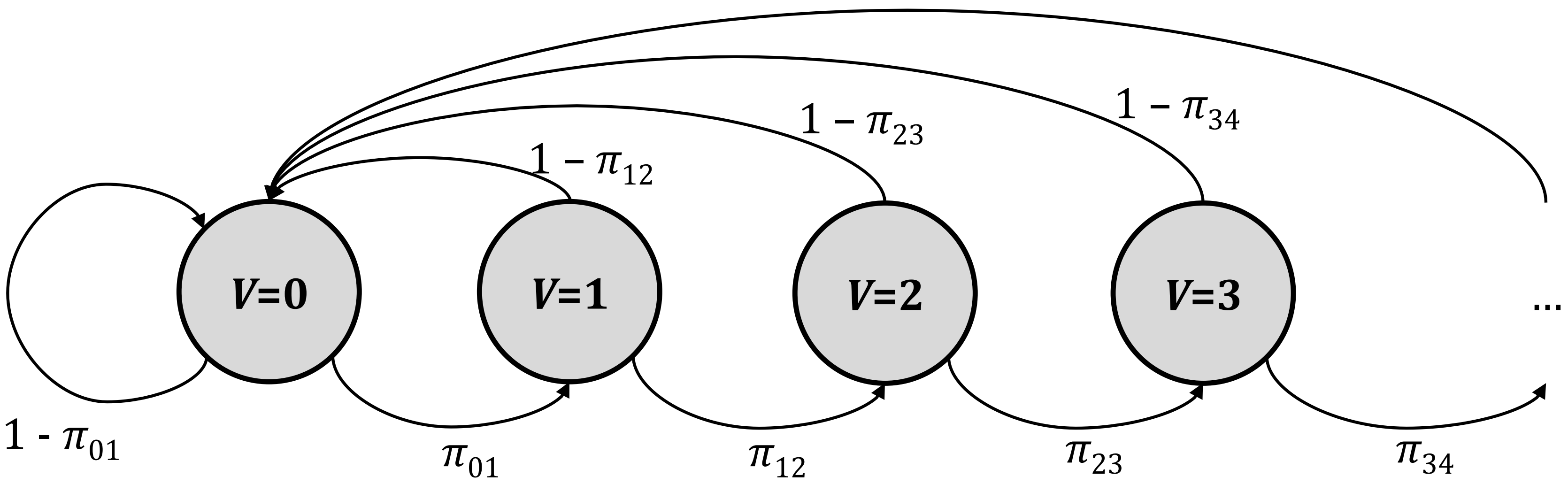
$$p_1(t) = \Pr(Y(t) = 1 \mid Y(0) = 1) = (1-\phi)\exp\left(-\frac{\zeta t}{\phi(1-\phi)}\right) + \phi$$

Model for PIR data

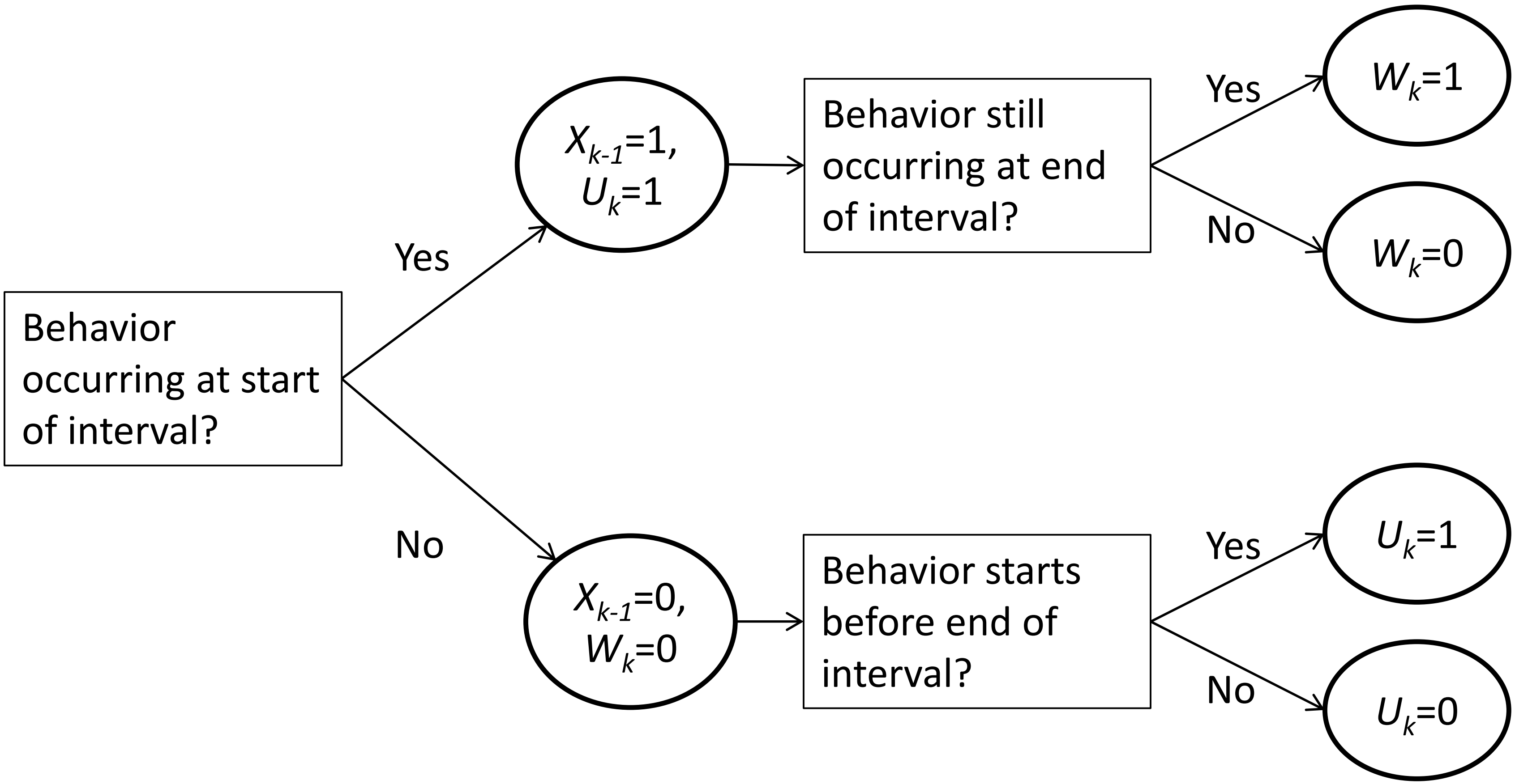
- Define V_k as the number of consecutive intervals where behavior is present:

$$V_k = k - \max\left\{0 \leq j \leq k : U_j = 0\right\}.$$

- Under the Alternating Poisson Process, $V_1,\ldots,V_K$ follow a DTMC on the space {0,1,2,3,…}.



Augmented interval recording (AIR)



Model for AIR data

- Define $Z_k = U_k + W_k + X_k$.
- Under the alternating Poisson process, $Z_1,\ldots,Z_{K/2}$ follow a DTMC on $\{0,1,2,3\}$, with transition probabilities $\pi_{ab}= \Pr(Z_k = b \mid X_{k\text{-}1} = a)$