Some Markov models for direct observation of behavior

James E. Pustejovsky
Northwestern University

May 29, 2013

Direct observation of behavior

- Quantities of interest
  - **Prevalence**: proportion of time that a behavior occurs
  - **Incidence**: rate at which new behavioral events begin
  - Intensity, contingency, others

- Applications in psychology and education research
  - Measurement of teaching practice
  - Measurement of student behavior
  - Evaluating interventions for individuals with disabilities
  - Other examples in animal behavior, organizational psychology, social work, exercise physiology

Observation recording methods

- How to turn direct observation of a "behavior stream" into data?

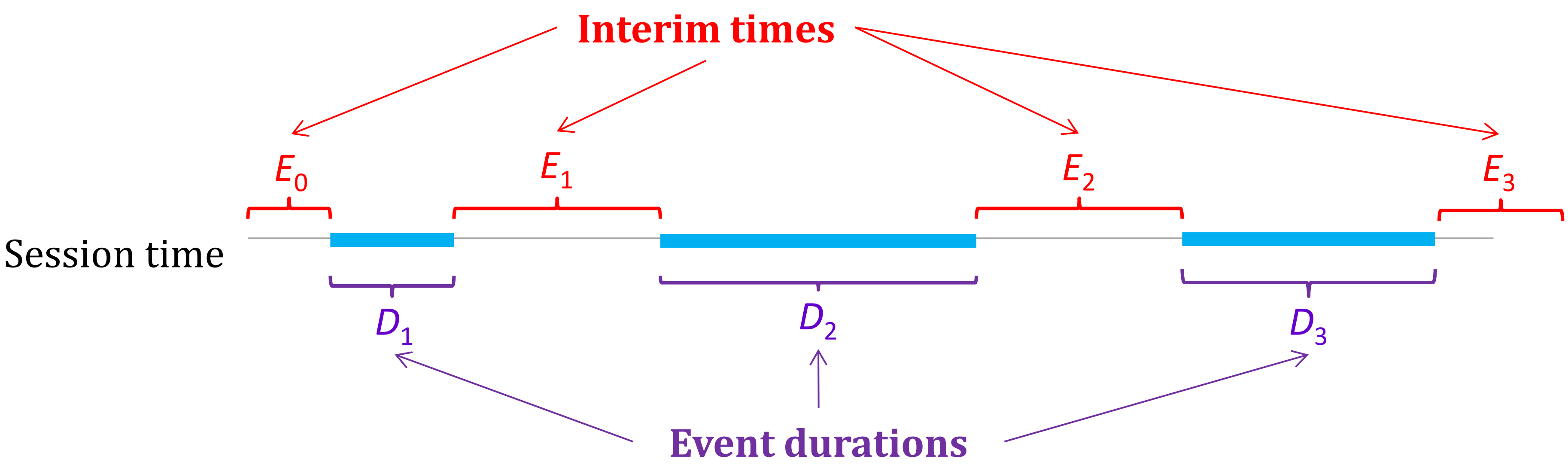
- Continuous recording methods
  - Produce rich data, amenable to sophisticated modeling
  - Effort-intensive

- Discontinuous recording methods
  - Less demanding methods needed in field settings
  - Momentary time sampling
  - Interval recording
  - Other possibilities?

Outline

- Model for behavior stream as observed
- Momentary time sampling
- Interval recording
- Some novel proposals
- Efficiency considerations

A model for the behavior stream



$$Y(t) = \begin{cases} 0 & \text{behavior is not occuring at time } t \\ 1 & \text{behavior is occuring at time } t \end{cases}$$

Alternating Poisson Process

Assumptions

1. Event durations: $D_j \sim \text{Exp}(1/\mu), \quad j = 1, 2, 3, \ldots$
2. Interim times: $E_j \sim \text{Exp}(1/\lambda), \quad j = 1, 2, 3, \ldots$
3. Event durations and interim times are all mutually independent.
4. Process is in equilibrium.

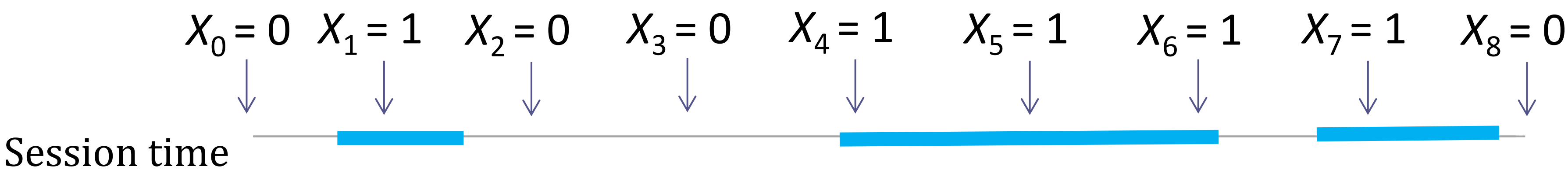
Under this model:

- Prevalence $\phi = \mu / (\mu + \lambda)$
- Incidence $\zeta = 1 / (\mu + \lambda)$

Momentary time sampling (MTS)

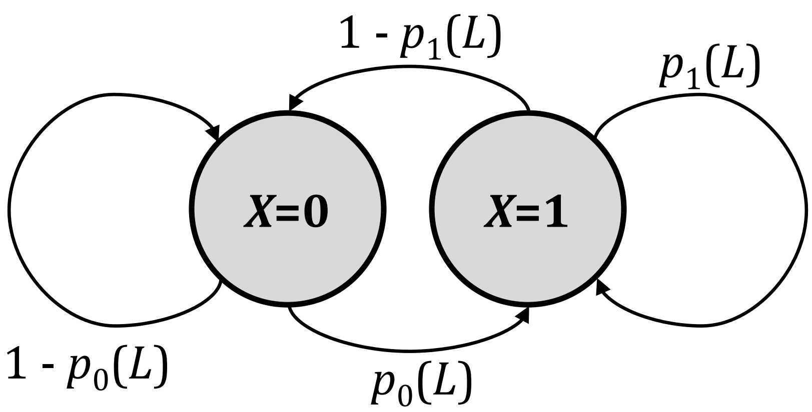
- $(K + 1)$ moments, equally spaced at intervals of length L.
- Observer records the presence or absence of a behavior at each moment
- Recorded data are

$$X_k = Y(kL), \quad k = 0, \ldots, K$$



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).



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

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

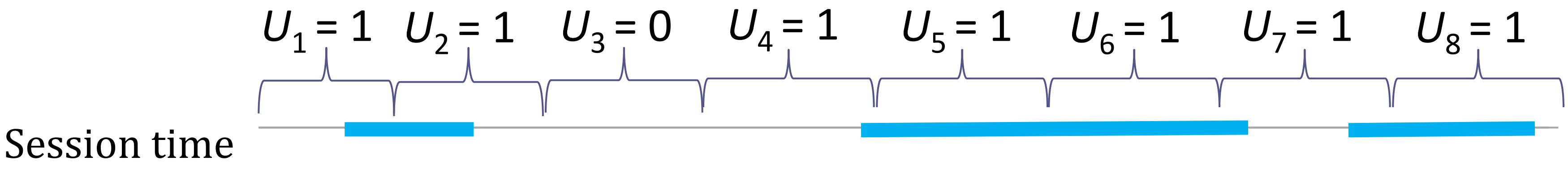
MTS model, continued

- Maximum likelihood estimators of ϕ and ζ have closed form expressions (Brown, Solomon, & Stephens, 1975).
- But $E\left(\bar{X}\right) = \phi$ under more general models.
- Extensive literature, lots of generalizations
  - stopping rules for observation time (Brown, Solomon, & Stephens, 1977, 1979; Griffin & Adams, 1983)
  - Irregular observation times (e.g., Cook, 1999)
  - Random effects to describe variation across subjects (e.g., Cook et al., 1999)

Partial interval recording (PIR)

- Divide period into K intervals, each of length L.
- For each interval, observer records whether behavior occurred ***at any point*** during the interval.
- Recorded data are

$$U_k = I\left[0 < \int_{[0,L)} Y\left(\left(k-1\right)L+t\right)dt\right], \quad k = 1,\ldots,K.$$



PIR, continued

- Unlike MTS, the mean of PIR data is not readily interpretable:

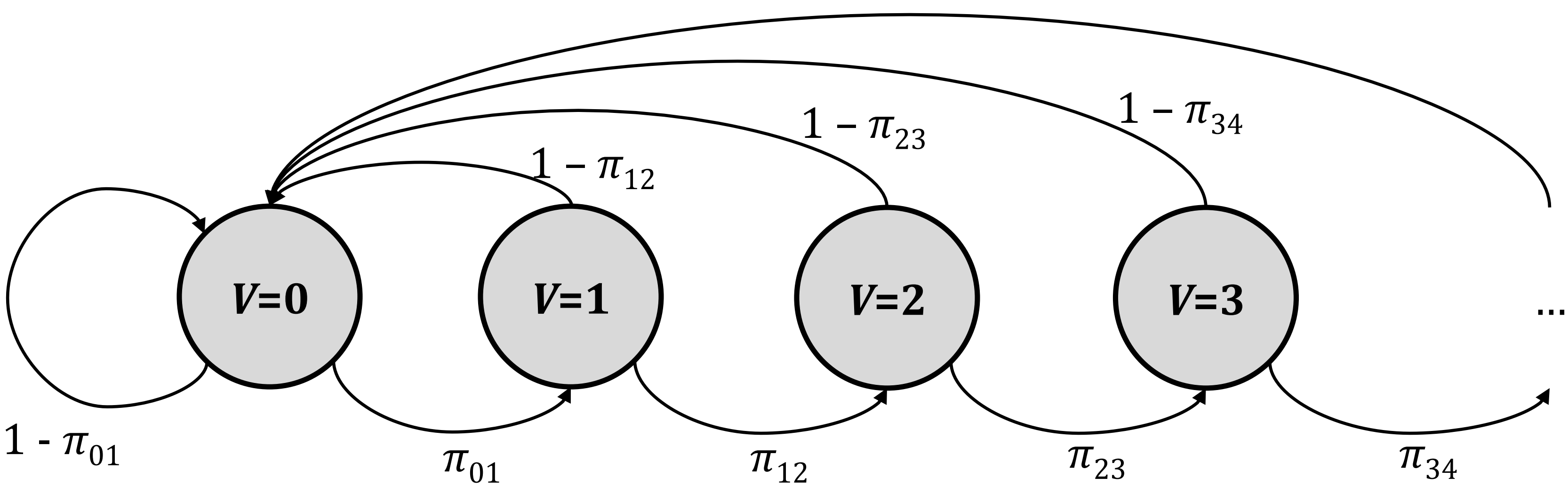
$$E\left(\bar{U}\right) = \phi + (1 - \phi)\left[1 - \exp\left(\frac{-\zeta L}{1 - \phi}\right)\right]$$

Model for PIR data

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

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

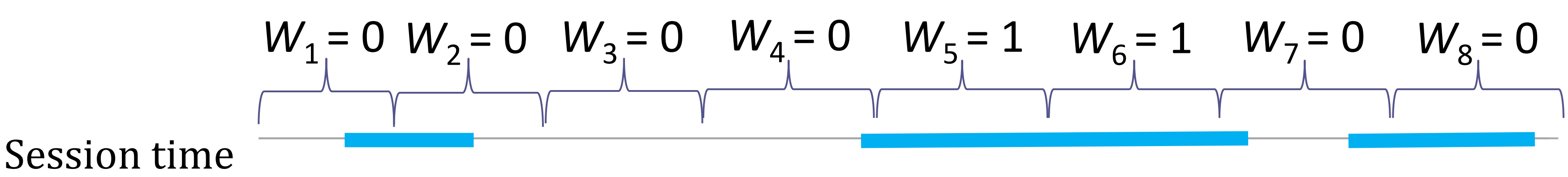
- Under the alternating Poisson process, $V_1,\ldots,V_K$ follow a DTMC on the space {0,1,2,3,…}.



Whole interval recording (WIR)

- Divide period into K intervals, each of length L.
- For each interval, observer records whether behavior occurred ***for the duration*** of the interval.
- Recorded data are

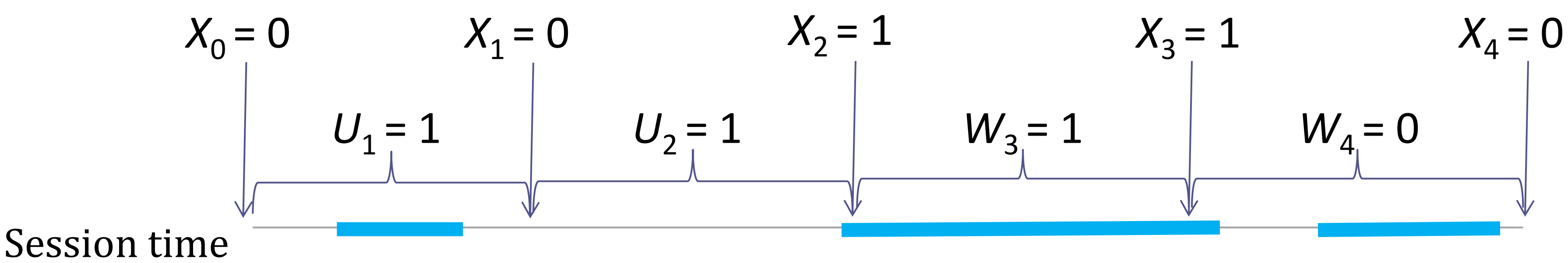
$$W_k = I\left[L = \int_{[0,L)} Y\left(\left(k-1\right)L + t\right)dt\right], \quad k = 1,\ldots,K.$$



- Equivalent to PIR for absence of event.

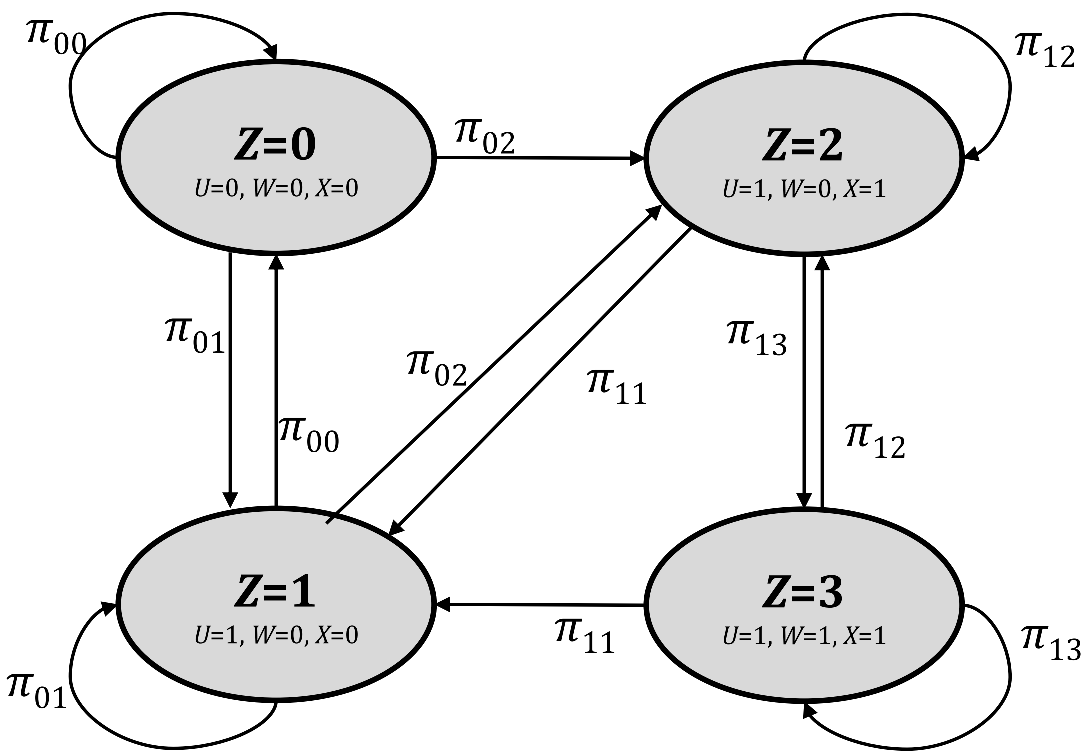
Augmented interval recording (AIR)

- Divide period into $K/2$ intervals, each of length $2L$.
- Use MTS at the beginning of each interval, to record $X_{k\text{-}1}$.
- If $X_{k\text{-}1} = 0$, use PIR for the remainder of the interval.
- If $X_{k\text{-}1} = 1$, use WIR for the remainder of the interval.



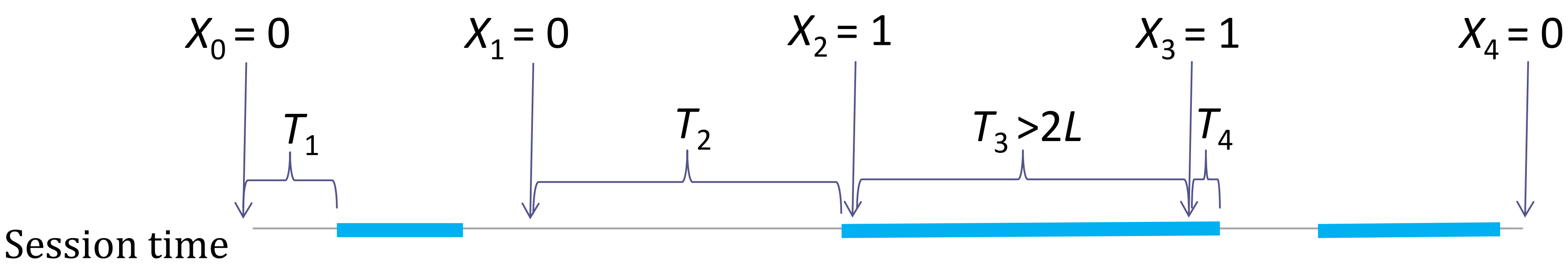
Model for AIR data

- 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 $\pi_{ab} = \Pr(Z_k = b \mid X_{k-1} = a)$



Intermittent transition recording (ITR)

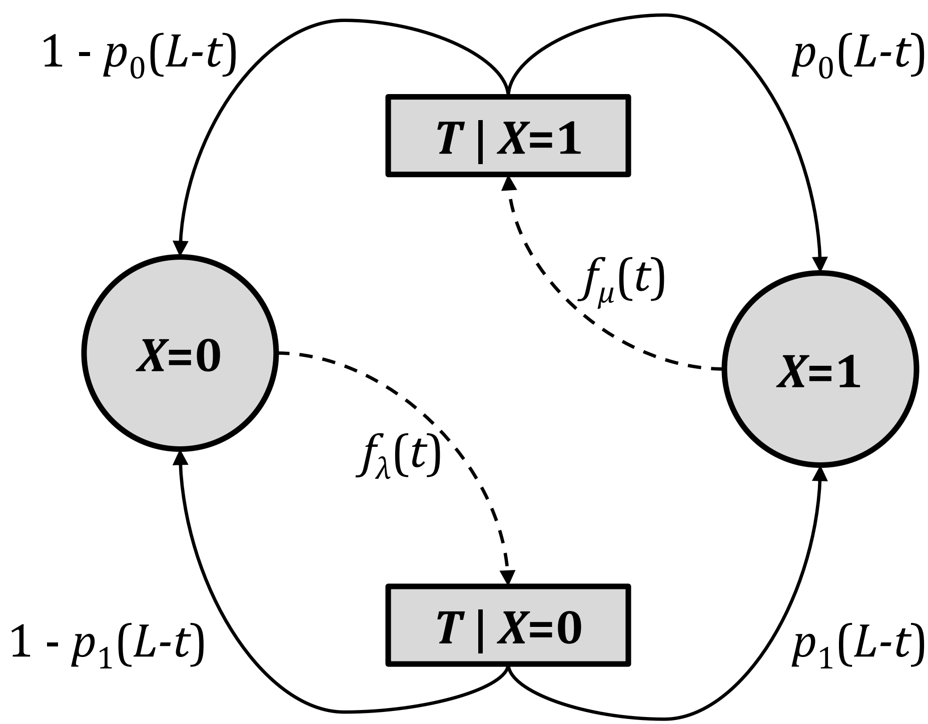
- Divide period into $K/2$ intervals, each of length $2L$.
- Use MTS at the beginning of each interval, to record X_{k-1}.
- Record time until next transition as T_k.



Model for ITR data

- Under the alternating Poisson process, $T_1, X_1, \ldots, T_{K/2}, X_{K/2}$ have the property that

$$F(T_k, X_k \mid T_1, X_1, \ldots, T_{k-1}, X_{k-1}) = F(T_k, X_k \mid X_{k-1})$$



Asymptotic relative efficiency

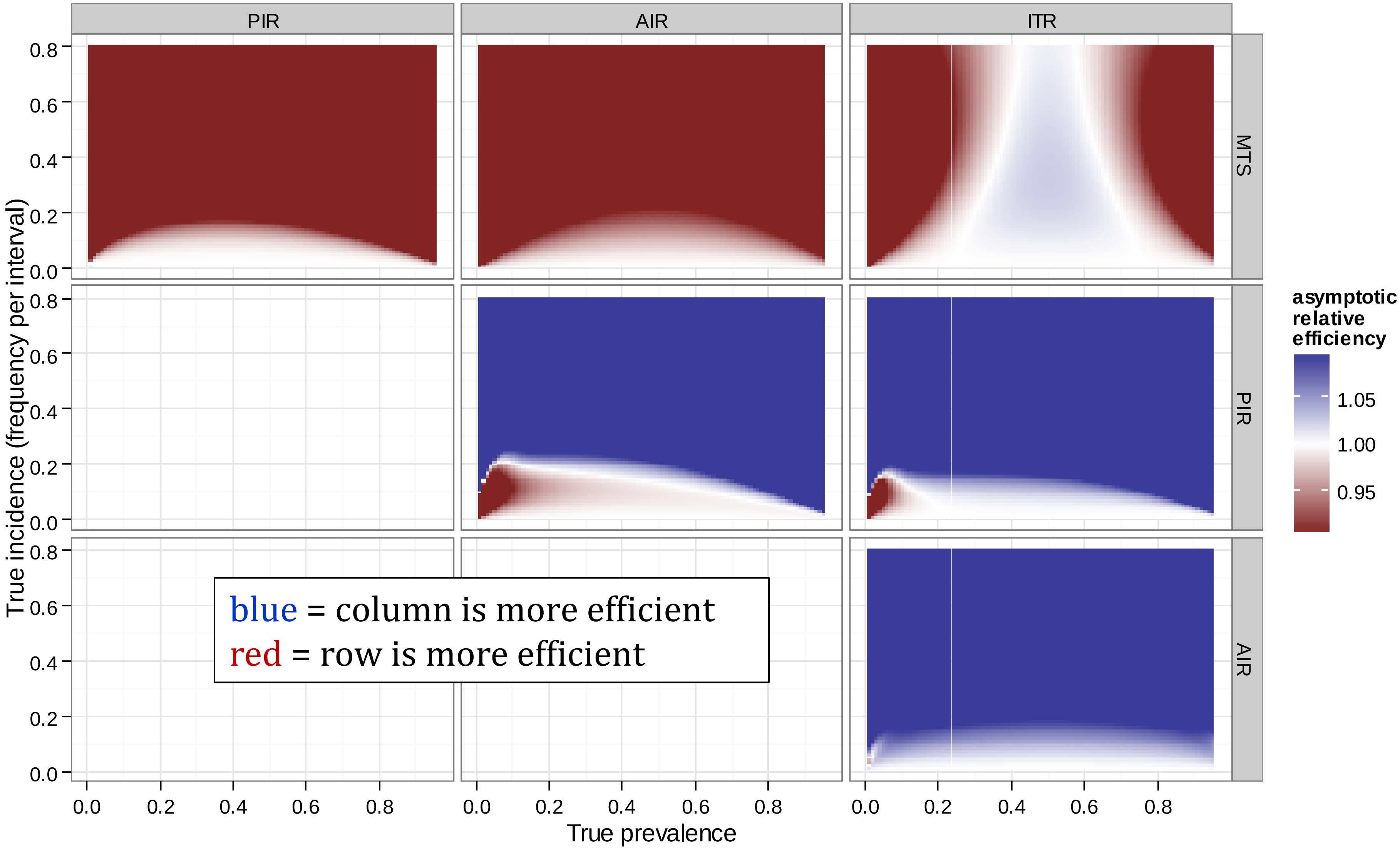
- Procedure $p, q \in$ {MTS, PIR, AIR, ITR}
- $\hat{\phi}_p, \hat{\zeta}_p$ are maximum likelihood estimators based on procedure *p*
- $V\left(\hat{\phi}_p\right), V\left(\hat{\zeta}_p\right)$ are asymptotic variances based on inverse of expected information matrix.

- Asymptotic relative efficiency of *p* versus *q*

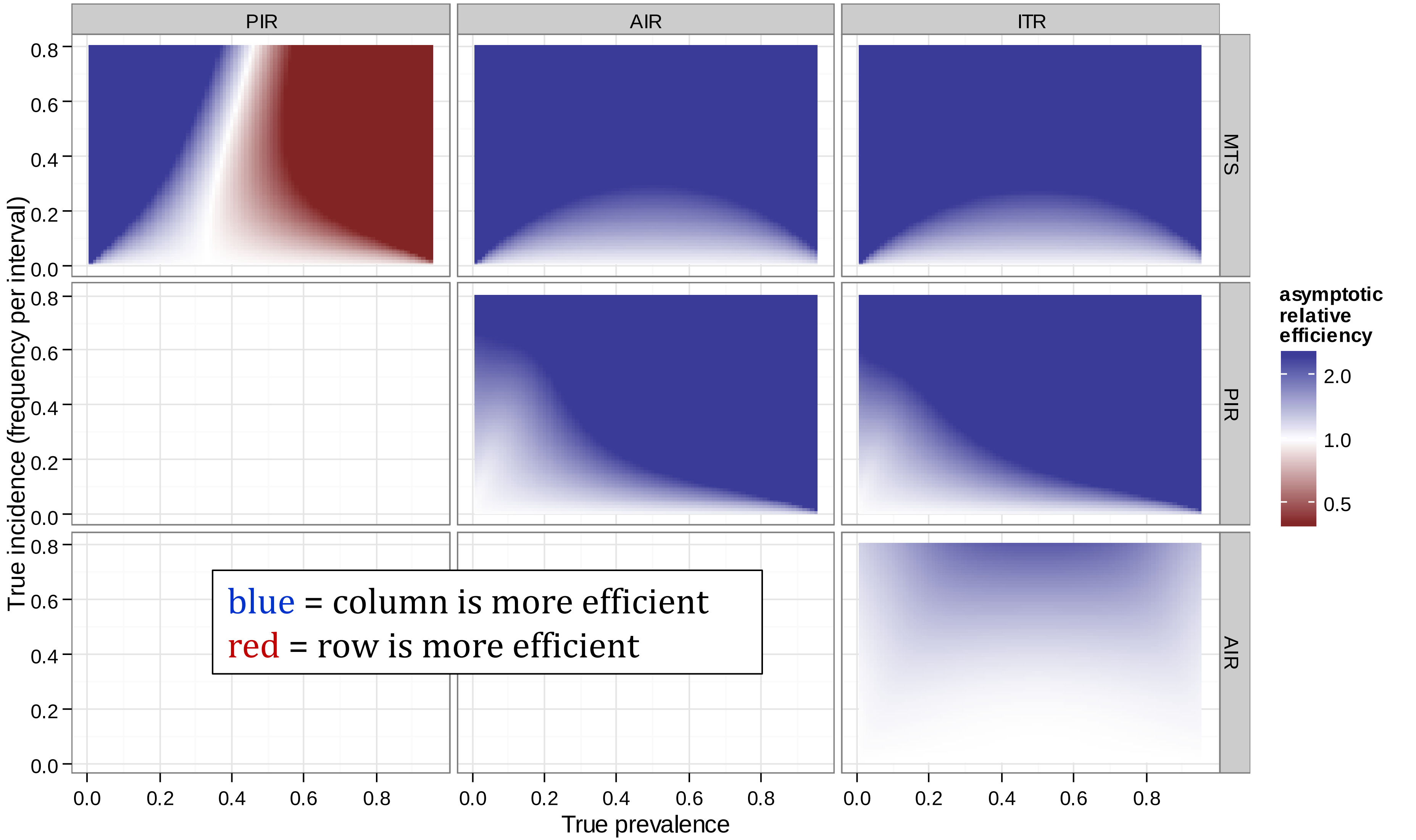
$$\text{ARE}\left(\hat{\phi}_p, \hat{\phi}_q\right) = \frac{\text{V}\left(\hat{\phi}_q\right)}{\text{V}\left(\hat{\phi}_p\right)}$$

$$\text{ARE}\left(\hat{\zeta}_p, \hat{\zeta}_q\right) = \frac{\text{V}\left(\hat{\zeta}_q\right)}{\text{V}\left(\hat{\zeta}_p\right)}$$

Asymptotic relative efficiency: Prevalence



Asymptotic relative efficiency: Incidence



Future work

- Evaluating these models & methods
  - Field testing
  - When is it okay to treat ML estimates from individual sessions as "pre-processing"?

- Lots still to do
  - Build data-collection software
  - Extensions to between-period regression models
  - Random period/subject effects
  - PIR, AIR, ITR under other distributional assumptions?

Questions? Comments?

pusto@u.northwestern.edu

PIR model

- Transition probabilities are uglier than MTS:

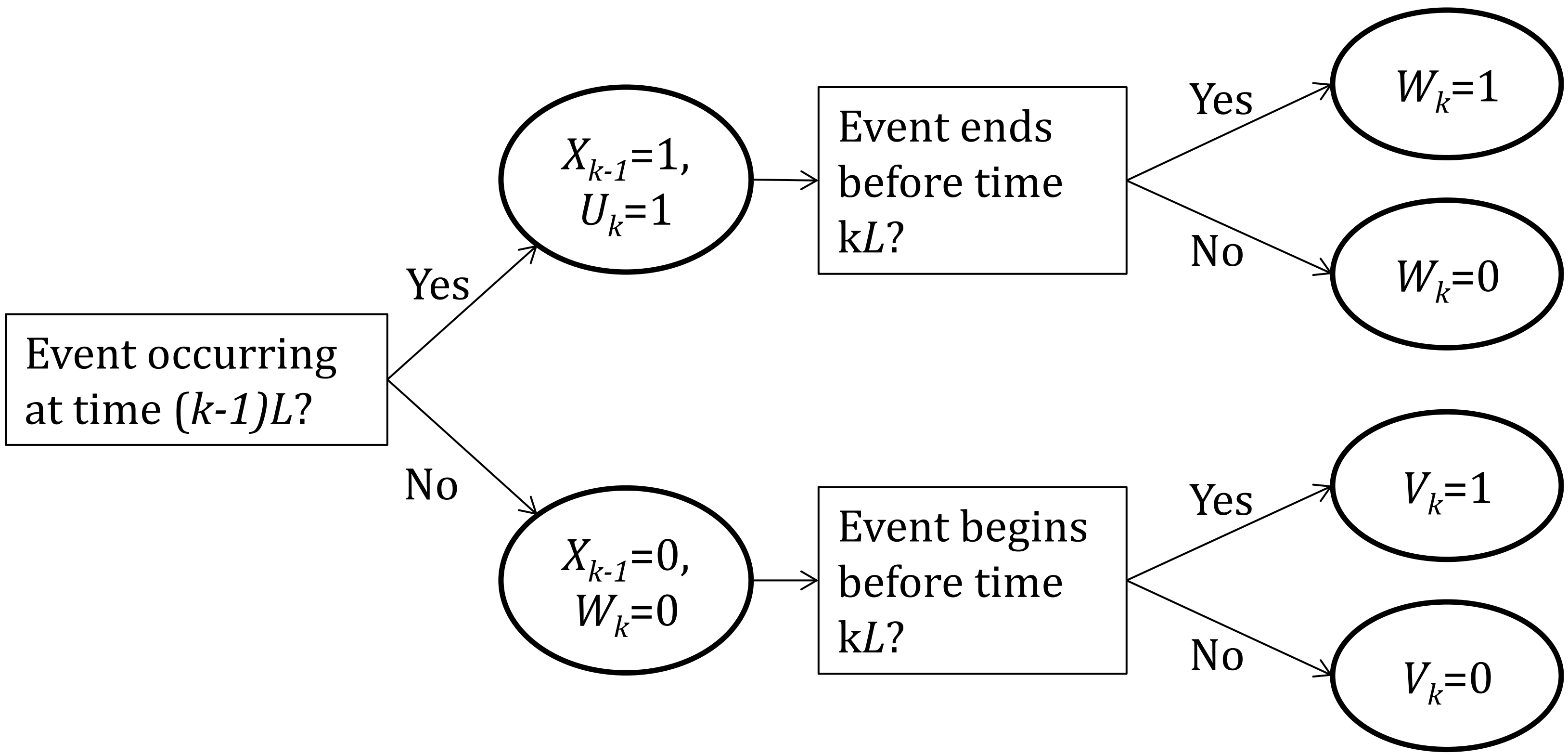
$$\pi_{j,j+1} = \Pr\left(V_k = j+1 \mid V_{k-1} = j\right) = 1 - e^{-L/\lambda}\left[1 - f^{(j)}(0)\right]$$

where

$$f(q) = \frac{\phi - (\phi - q)e^{-L\zeta/[\phi(1-\phi)]}}{1-(1-q)e^{-L\zeta/(1-\phi)}}$$

and $f^{(j)}$ is the j-fold recursion of f.

AIR, continued



AIR model

- Transition probabilities are given by

$$\pi_{00} = e^{-2L/\lambda}$$

$$\pi_{10} = 0$$

$$\pi_{01} = 1 - e^{-2L/\lambda} - p_0(2L)$$

$$\pi_{11} = \frac{\lambda}{\mu} p_0(2L)$$

$$\pi_{02} = p_0(2L)$$

$$\pi_{12} = 1 - e^{-2L/\lambda} - \frac{\lambda}{\mu} p_0(2L)$$

$$\pi_{03} = 0$$

$$\pi_{13} = e^{-2L/\lambda}$$

ITR model

- Under alternating Poisson process,

$$f\left(t_1, x_1, \ldots, t_{K/2}, x_{K/2} \mid x_0\right) = \prod_{k=1}^{K/2} f\left(t_k \mid x_{k-1}; \mu, \lambda\right) f\left(x_k \mid t_k, x_{k-1}; \mu, \lambda\right)$$