Modeling Criminal Careers as Departures From a Unimodal Population Age-Crime Curve: The Case of Marijuana Use

(2012)

Donatello Telesca, Elena A. Erosheva,
Derek A. Kreader, & Ross Matsueda

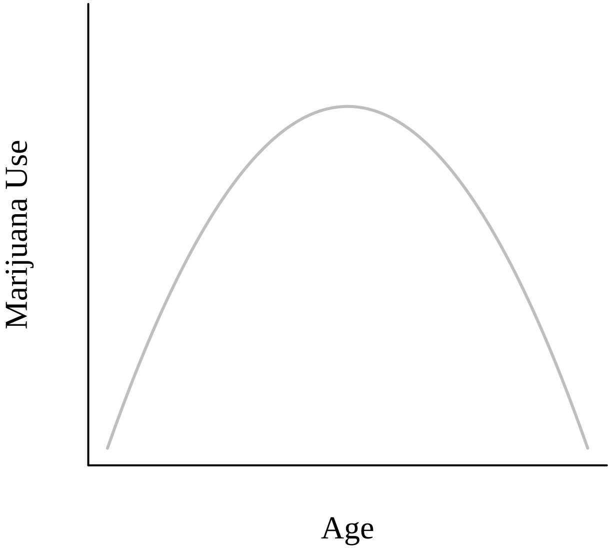
April 15, 2014

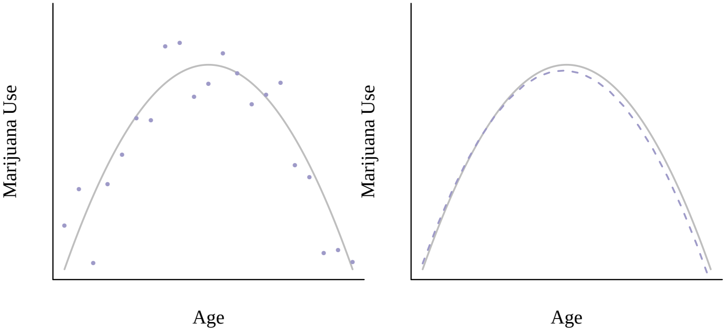
- Telesca et al. (2012) extends Telesca and Inoue (2008) by applying the methods of Telesca and Inoue (2008) to a study of individual age-crime curves in the context of marijuana use over the course of young adulthood

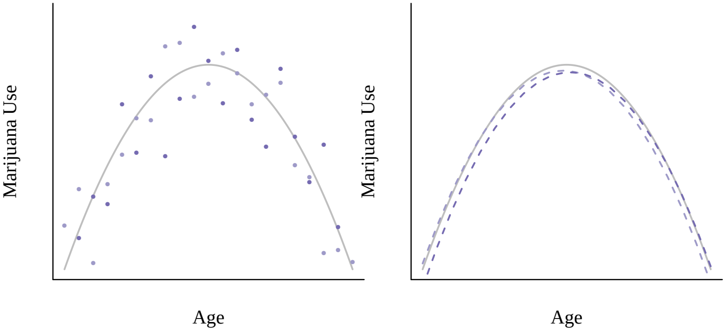
- Representative of criminal career trajectories
- Used to study:
  - Age of onset of offending
  - Timing of peak offending
  - Duration of criminal careers
  - "Types" of criminal careers
  - Other features of criminal careers
- When combined with individual characteristics, can also be used to describe relationship between individual characteristics and features of criminal careers

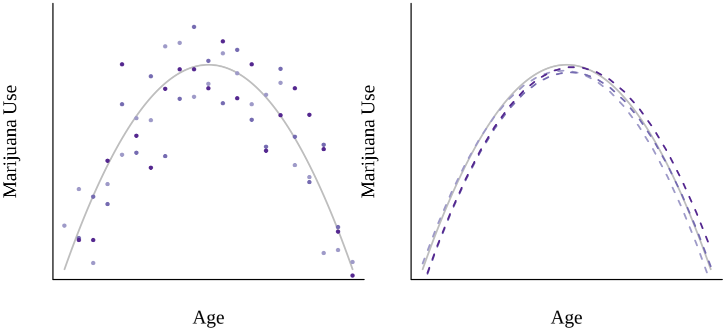
- Relatively new field of statistics
- The same/similar methods referred to as curve registration, warping regression models, structural averaging ...
- Useful when:
  - Data is observed as a time series for many individuals
  - Features of the time-outcome relationship (peaks, inflection points) are of interest
  - Substantial variation in individual curves

What would we need the data to look like for these questions to be easy to answer?

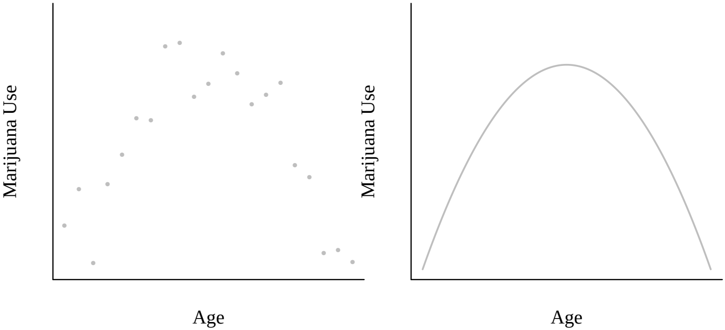

Marijuana Use
Age

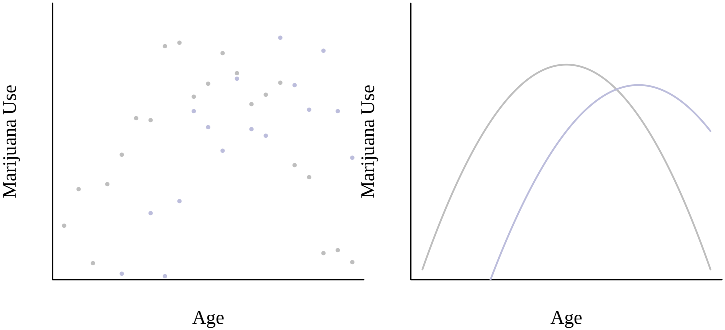

Marijuana Use
Age
Marijuana Use
Age

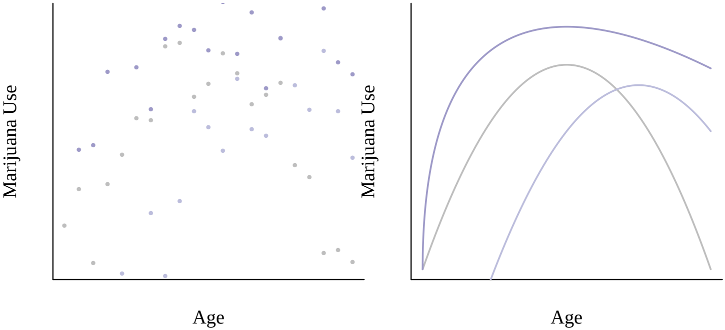

Marijuana Use
Age
Marijuana Use
Age

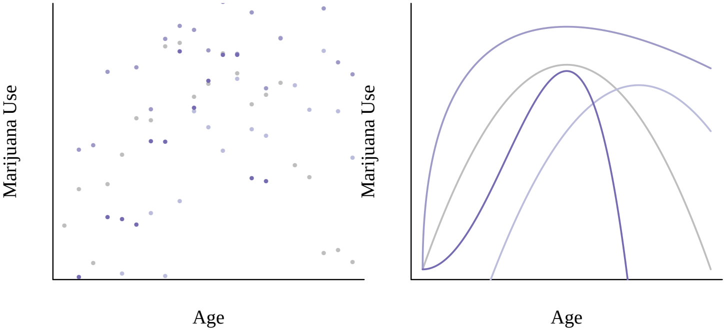

Marijuana Use
Age
Marijuana Use
Age

What does the data actually look like?

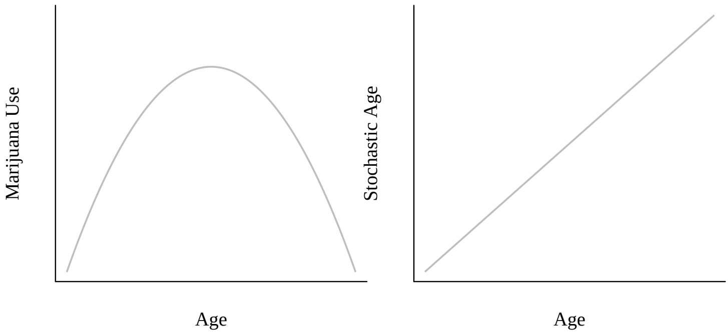

Marijuana Use
Age
Marijuana Use
Age

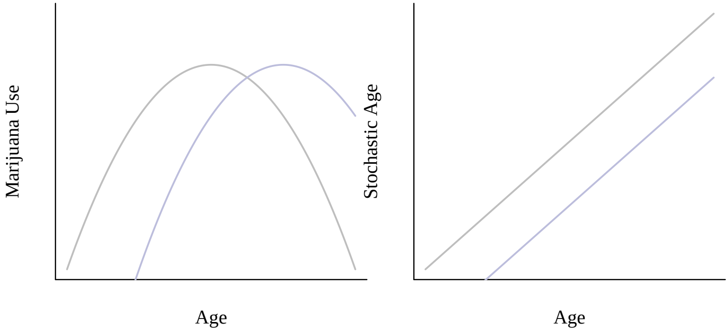

Marijuana Use
Age
Marijuana Use
Age

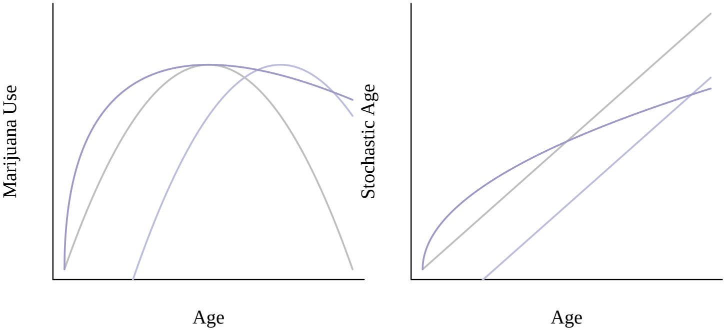

Marijuana Use
Age
Marijuana Use
Age

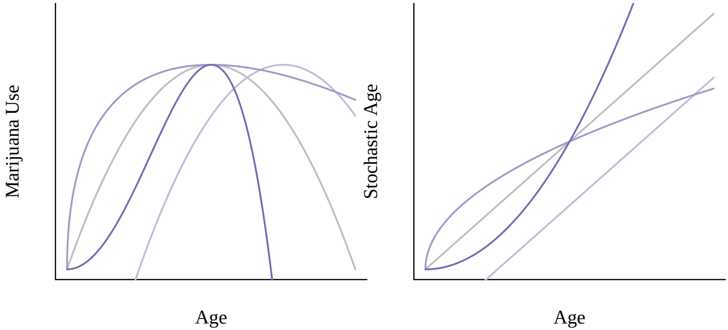

Marijuana Use
Age
Marijuana Use
Age

How does the functional data analysis framework allow us to characterize data of this kind?


Marijuana Use
Age
Stochastic Age
Age


Marijuana Use
Age
Stochastic Age
Age

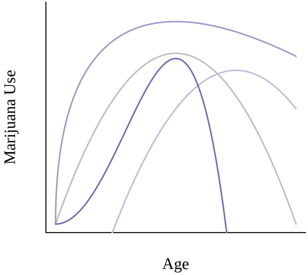

Marijuana Use
Age
Stochastic Age
Age


Marijuana Use
Age
Stochastic Age
Age

- Formulation used by Telesca et al. (2012) introduced in Ramsay and Li (1998)
- Introduces model where
    - All individuals share a stochastic time-response function
    - Variation in observed time-response curves is generated by individual, strictly monotonic functions mapping time to stochastic time, called "warping functions" and noise
        - Flexibility in modeling warping functions obtained by using B-spline basis functions
        - Smoothness of estimated warping functions obtained by penalization of relative curvature

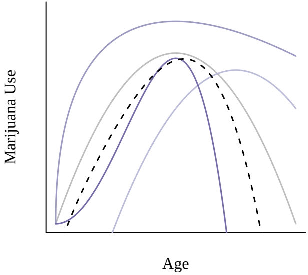
- Telesca and Inoue (2008) developed a Bayesian framework for functional data analysis based on Ramsay and Li (1998)
    - Allows information to be shared across curves
    - Introduces a temporal misalignment window
    - Incoporates smoothing penalty directly into the model
    - Gives a formal account of amplitude and phase variability
    - Allows for exact inference

- Telesca etal. (2012) is an extension of Telesca and Inoue (2008)
  - Extends the curve registration models, developed for continuous data, to count data
  - Allows scale and time-transformations to depend on individual covariates



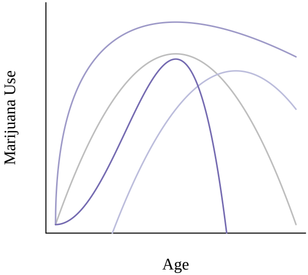
- Model mean trajectory as a degree two polynomial in time/age
- Incorporated individual variation using random effects and/or latent classes

- **Random Effects**: Raudenbush and Chan (1993)
- **Latent Classes**: Nagin and Land (1993); Roeder et al. (1999)
- **Random Effects & Latent Classes**: Muthén and Shedden (1999)



- Model mean trajectory as a degree two polynomial in time/age
- Incorporate individual variation using random effects and/or latent classes

- **Random Effects**: Raudenbush and Chan (1993)
- **Latent Classes**: Nagin and Land (1993); Roeder et al. (1999)
- **Random Effects & Latent Classes**: Muthén and Shedden (1999)



$$Y_{ij}|\lambda_i\left(t_j, \boldsymbol{X}_i\right) \sim \mathcal{P}\left(\lambda_i\left(t_j, \boldsymbol{X}_i\right)\right)$$

$$\begin{aligned}\lambda_i\left(t_j, \boldsymbol{X}_i\right) &= \underbrace{a_i\left(\boldsymbol{X}_i\right)}_{\text{Amplitude}} \underbrace{S\left(t_j, \boldsymbol{\beta}\right)}_{\text{Shape}} \circ \underbrace{\mu_i\left(t_j, \boldsymbol{\phi}_i; \boldsymbol{X}_i\right)}_{\text{Phase Shift}} \\ &= a_i\left(\boldsymbol{X}_i\right) S\left(\mu_i\left(t_j, \boldsymbol{\phi}_i; \boldsymbol{X}_i\right), \boldsymbol{\beta}\right)\end{aligned}$$

$a_i\left(\boldsymbol{X}_i\right)$

- $\mathbb{E}\left[a_i|\boldsymbol{X}_i\right] = \exp\left\{\boldsymbol{X}_i'\boldsymbol{b}_a\right\}$
- Gamma conditional prior distribution of $a_i|\boldsymbol{X}_i, \boldsymbol{b}_a$
    - Conjugate prior distribution for Poisson rate parameter
    - Naturally accommodates over dispersion

$S(t_j, \boldsymbol{\beta})$

- $S(t_j, \boldsymbol{\beta}) = \mathcal{S}_B(t_j)' \boldsymbol{\beta}$
  - $\mathcal{S}_B(t_j)$ set of K B-spline basis functions of order four evaluated at time t
- Positivity and unimodality of $S(t_j, \boldsymbol{\beta})$ imposed by restrictions on $\boldsymbol{\beta}$
- Second order shrinkage prior on a reparametrization of $\boldsymbol{\beta}$ shrinks $S(t_j, \boldsymbol{\beta})$ towards a piecewise linear trajectory

$\mu_i\left(t_j, \boldsymbol{\phi}_i; \boldsymbol{X}_i\right)$

- $\mu_i\left(t_j, \boldsymbol{\phi}_i; \boldsymbol{X}_i\right) = \mathcal{S}_\mu\left(t_j\right)' \boldsymbol{\phi}_i\left(\boldsymbol{X}_i\right)$
  - $\mathcal{S}_\mu\left(t_j\right)$ set of Q B-spline basis functions of order four evaluated at time t
- Monotonicity and boundedness of $\mu_i\left(t_j, \boldsymbol{\phi}_i; \boldsymbol{X}_i\right)$ imposed by restrictions on $\boldsymbol{\phi}_i$
- First order random walk shrinkage prior on a reparametrization of $\boldsymbol{\phi}_i$ shrinks $\boldsymbol{\phi}_i$ towards the B-spline coefficients that yield the identity function, $\mu_i\left(t_j, \boldsymbol{\phi}_i; \boldsymbol{X}_i\right) = t_j$
- $\mathbb{E}\left[\phi_{iq} | \boldsymbol{X}_i, \boldsymbol{b}_\phi\right] = \Gamma_q + \boldsymbol{X}_i' \boldsymbol{b}_\phi$

- Use Markov Chain Monte Carlo to fit model by combination of Gibbs sampling and Metropolis-Hastings sampling algorithms
- Use posterior predictive loss to choose number of interior knots

Model allows us to explicitly test:

- Relationship between level of offending and individual characteristics
- Unimodality of age-crime curve
- Relationship between phase shift and individual characteristics

- Better familiarity with relevant concepts in foundational papers:
  - Ramsay and Li (1998)
  - Lang and Brezger (2004)
  - Telesca and Inoue (2008)
- Work through derivations of:
  - Posterior predictive loss for Poisson case
  - Conditional posterior distributions of parameters
  - MCMC algorithm used
- Begin working through code

Lang, S. and A. Brezger (2004, March). Bayesian P-Splines. *Journal of Computational and Graphical Statistics 13*(1), 183–212.

Muthén, B. and K. Shedden (1999, June). Finite Mixture Modeling with Mixture Outcomes Using the EM Algorithm. *Biometrics 55*(June), 463–469.

Nagin, D. S. and K. C. Land (1993). Age, Criminal Careers, and Population Heterogeneity: Specification and Estimation of a Nonparametric, Mixed Poisson Model. *Criminology 31*(3), 327–362.

Ramsay, J. O. and X. Li (1998, May). Curve Registration. *Journal of the Royal Statistical Society: Series B 60*(2), 351–363.

Raudenbush, S. W. and W.-S. Chan (1993, December). Application of a Hierarchical Linear Model to the Study of Adolescent Deviance in an Overlapping Cohort Design. *Journal of Consulting and Clinical Psychology 61*(6), 941–951.

Roeder, K., K. G. Lynch, and D. S. Nagin (1999). Modeling Uncertainty in Latent Class Membership: A Case Study in Criminology. *Journal of the American Statistical Association 94*(447), 766–776.

Telesca, D., E. A. Erosheva, D. A. Kreager, and R. L. Matsueda (2012, December). Modeling Criminal Careers as Departures From a Unimodal Population AgeCrime Curve: The Case of Marijuana Use. *Journal of the American Statistical Association 107*(500), 1427–1440.

Telesca, D. and L. Y. T. Inoue (2008, March). Bayesian Hierarchical Curve Registration. *Journal of the American Statistical Association 103*(481), 328–339.