A Likelihood Ratio Approach for Detecting Behavioral Changes in Device Usage Over Time

Rachel Longjohn and Padhraic Smyth, University of California Irvine

Background

Evidence: categorical sequence of user-generated event data on a device

Ex: sending texts or emails, making calls, using apps

Forensic question: is there evidence of a change in event patterns, or changepoint, during this sequence?

Is there evidence of a changepoint here specifically? Old method (1)

Is there evidence of a changepoint anytime in this time window? This method

A changepoint in the sequence may indicate that the sequence of events was generated by two different individuals rather than a single individual.

Hypotheses: one source vs. two source

H_1	H_2
: there is no changepoint in the sequence; it was generated by a single individual	: there is one changepoint in the sequence; it was generated by two different individuals

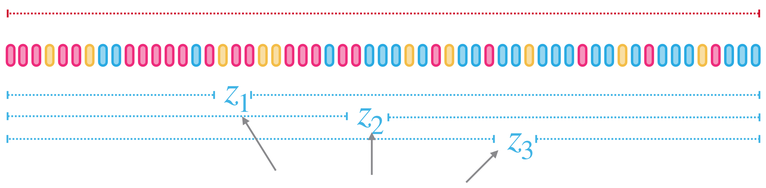
Methods

Likelihood ratio (Bayes factor):

$$\frac{P(H_1)}{P(H_2)} \times \frac{P(E \mid H_1)}{P(E \mid H_2)} = \frac{P(H_1 \mid E)}{P(H_2 \mid E)}$$

Model:

$X_{1:N} \mid H_1, \theta_1 \overset{iid}{\sim} Categ(\theta_1)$ → Count events per category → $C_{1:N}$

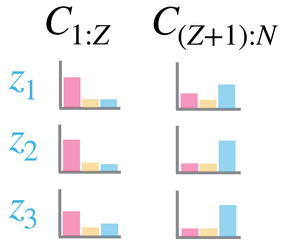


Relevant case information could be used to specify what changepoints are probable

$X_{1:Z} \mid H_2, \theta_1 \overset{iid}{\sim} Categ(\theta_1)$

$X_{(Z+1):N} \mid H_2, \theta_2 \overset{iid}{\sim} Categ(\theta_2)$

Count before and after each possible Z → $C_{1:Z}$, $C_{(Z+1):N}$

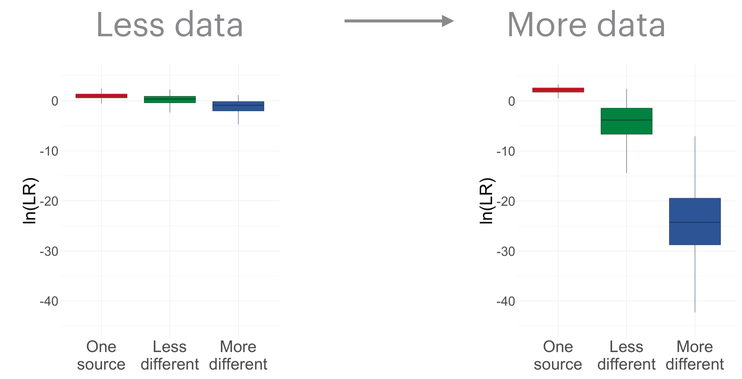


Formula for the likelihood ratio: conjugate prior for θ; easy to calculate

$$LR = \frac{B(\alpha + C_{1:N})B(\alpha)}{\sum_z B(\alpha + C_{1:z})B(\alpha + C_{(z+1):N})p(Z = z)}$$

Results + Discussion

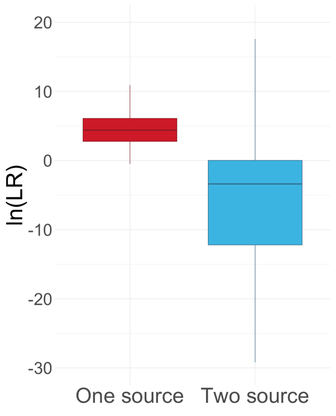
Simulations: vary amount of data and differences in event patterns



Experiments: evaluate performance using email dataset (2)

TPR@1	FPR@1	AUC
97.7%	25.6%	92.7%

C_{llr}	C_{llr}^{min}	C_{llr}^{cal}
1.13	0.37	0.75



(1) Longjohn, Rachel, Padhraic Smyth, and Hal Stern. "Likelihood Ratios for Categorical Evidence with Applications in Digital Evidence." *AAFS Annual Conference*, 2022.

(2) Paranjape, Ashwin, Austin R. Benson, and Jure Leskovec. "Motifs in temporal networks." *Proceedings of the tenth ACM international conference on web search and data mining*. 2017.

This work was funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) through Cooperative Agreements 70NANB15H176 and 70NANB20H019 between NIST and Iowa State University, which includes activities carried out at Carnegie Mellon University, Duke University, University of California Irvine, University of Virginia, West Virginia University, University of Pennsylvania, Swarthmore College, and University of Nebraska, Lincoln.