Assessing Student Conceptual Competencies using Bayesian Networks

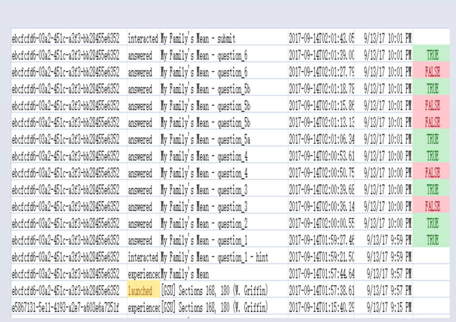
Wentao Yan Dr. Dennis Pearl Dr. Matthew Beckman

Department of Statistics, Penn State University

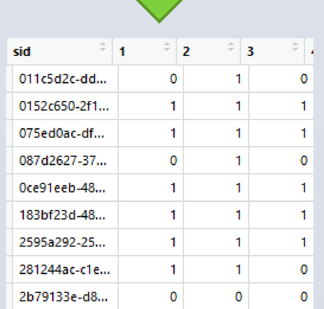
ABSTRACT

Web applications are often used in labs or online assignments to help students explore statistical principles. Examining web log files allows evaluation of student interaction with these apps to judge their conceptual competency. This poster will demonstrate an education assessment system aimed at giving instructors or individual students a quantitative measurement of students' competency in course concepts. We achieve this by building Bayesian Networks to model causal relationships between conceptual competency (θ), associated tasks response (x) and attributes of the individual (z) in the assessment. We applied our models on data for 75 introductory statistics students using interactive songs in project SMILES (www.causeweb.org/smiles). We analyzed the interaction of students with those prompts to judge conceptual knowledge before listening to the song.

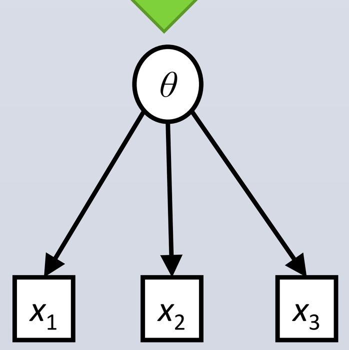
OBJECTIVES & DATA PIPELINE



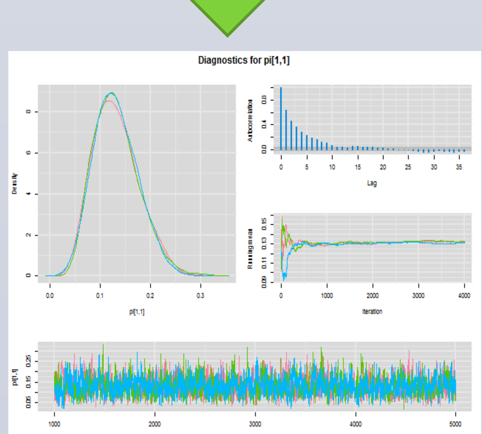
Generate a log file which contains students performance records in the educational App



Extract data from the log file and transform it into the model input format

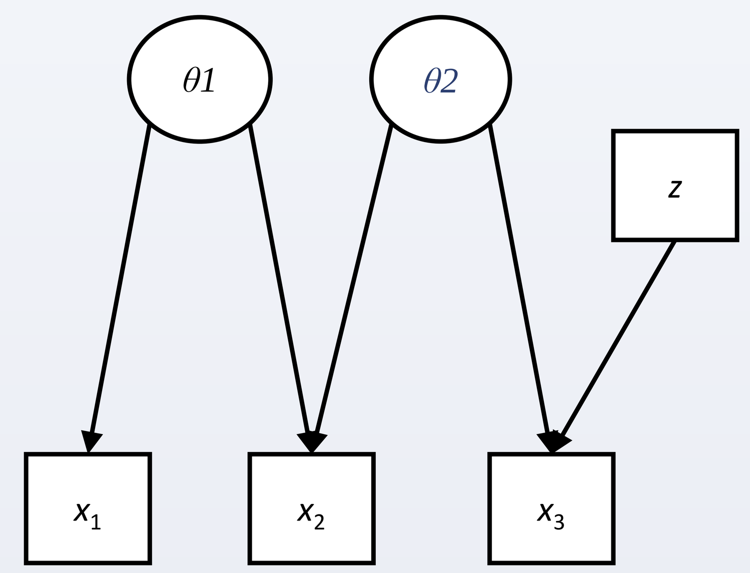


Build an evidence model based on the App question designs to associate conceptual competency (θ) with tasks response (x) and attributes of individuals (z)



Apply the model on pre-processed data and make posterior inference of latent variables.

METHODS



Θ = Conceptual Competency

ϑ ~ Bernoulli ($\boldsymbol{\lambda}$)

2 Levels of θ (Low, High)

$\boldsymbol{\lambda}$ means probability for High

x = Responses of tasks that measure concept

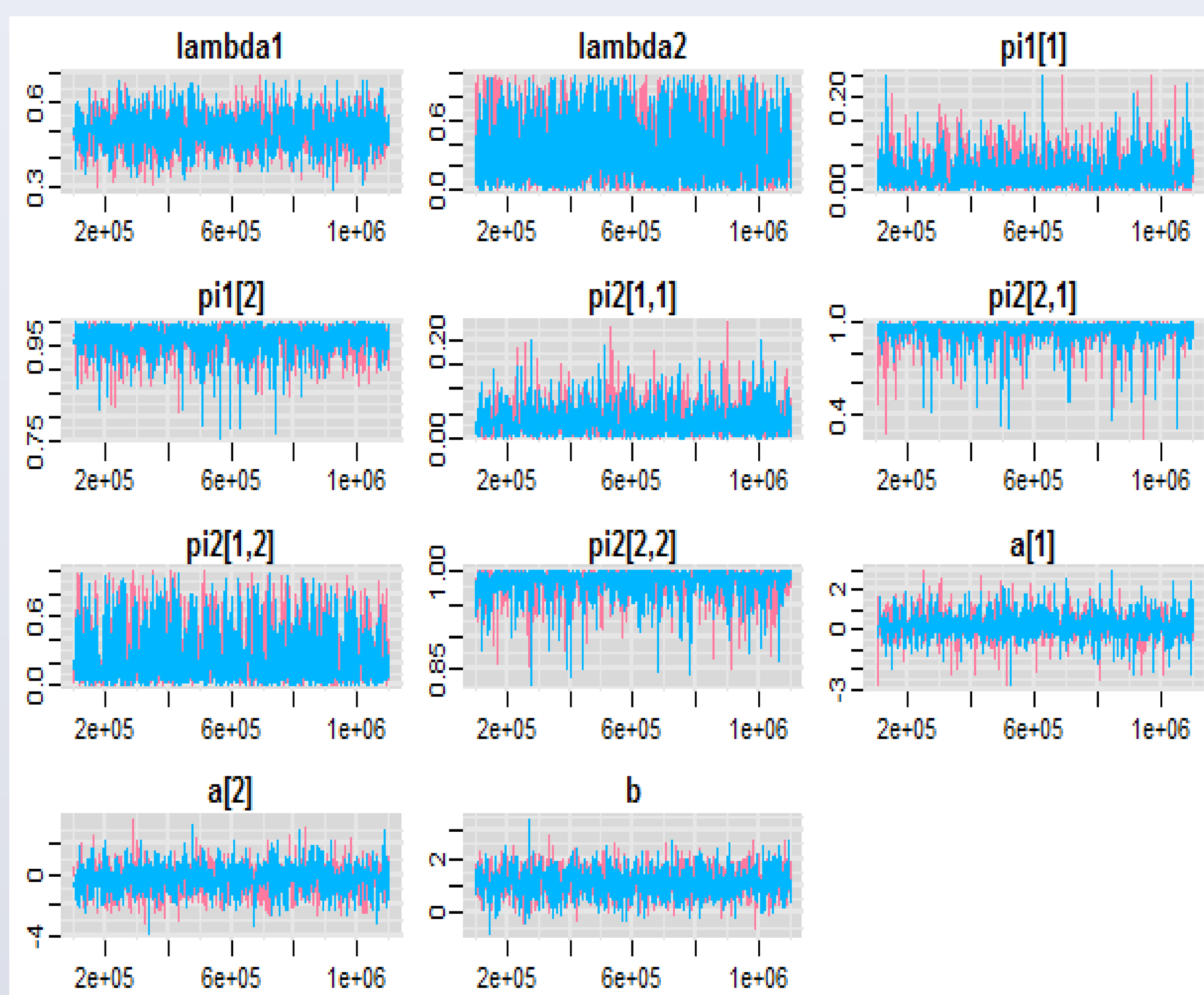
$(x_j \mid \vartheta = c)$ ~ Bernoulli ($\Pi\, j_{[c]}$)

$\Pi\, j_{[c]}$ is the probability of correct response on task j given $\vartheta = c$ and attribute z

$$p(\theta \mid x) \propto p(x \mid \theta)\, p(\theta)$$

z = Observable attributes of individuals

$\Pi\, j_{[c]}$ = sigmoid (a[c] + b * z)

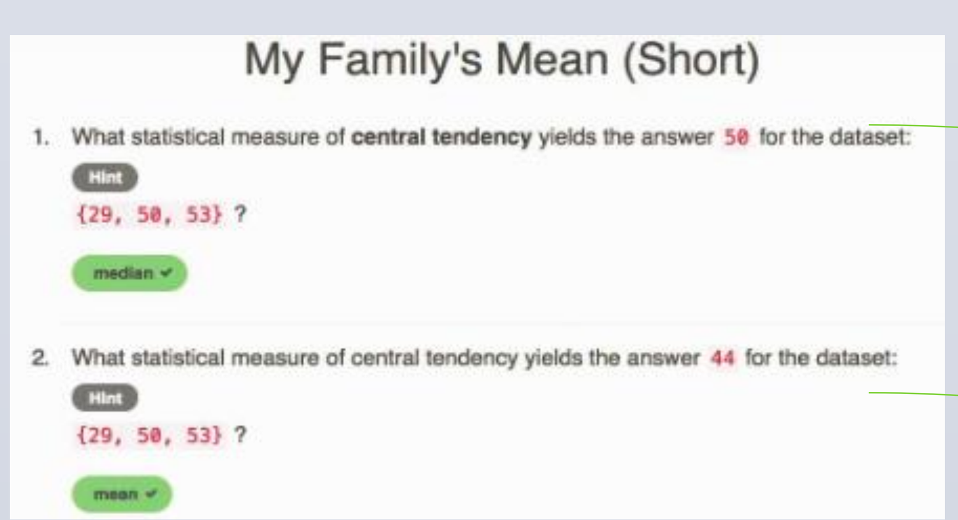


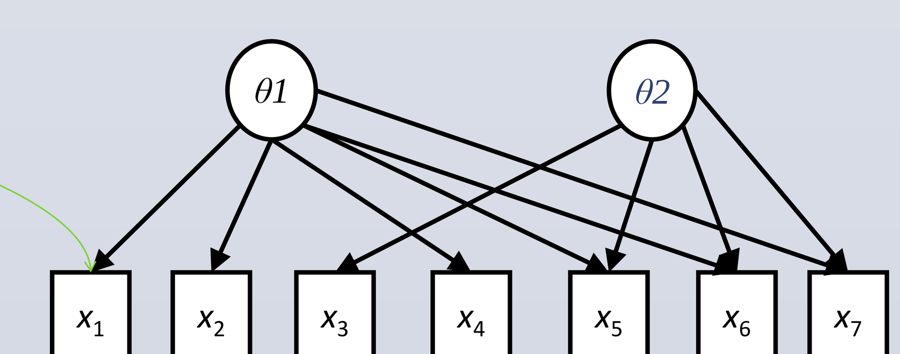
We built Bayesian Networks to model the relationship between conceptual competency (θ) , associated tasks response (**x**) and attributes of individuals (**z**) in the assessment. We could clearly interpret all the inference of parameters from fitting data, which is critical in education assessment.

Our model is fit using an MCMC algorithm to make inferences about the posterior distribution of latent variables. We ran two chains (colored in red & blue) independently on two computing cores concurrently to ensure that convergence happened in both cases in an efficient way.

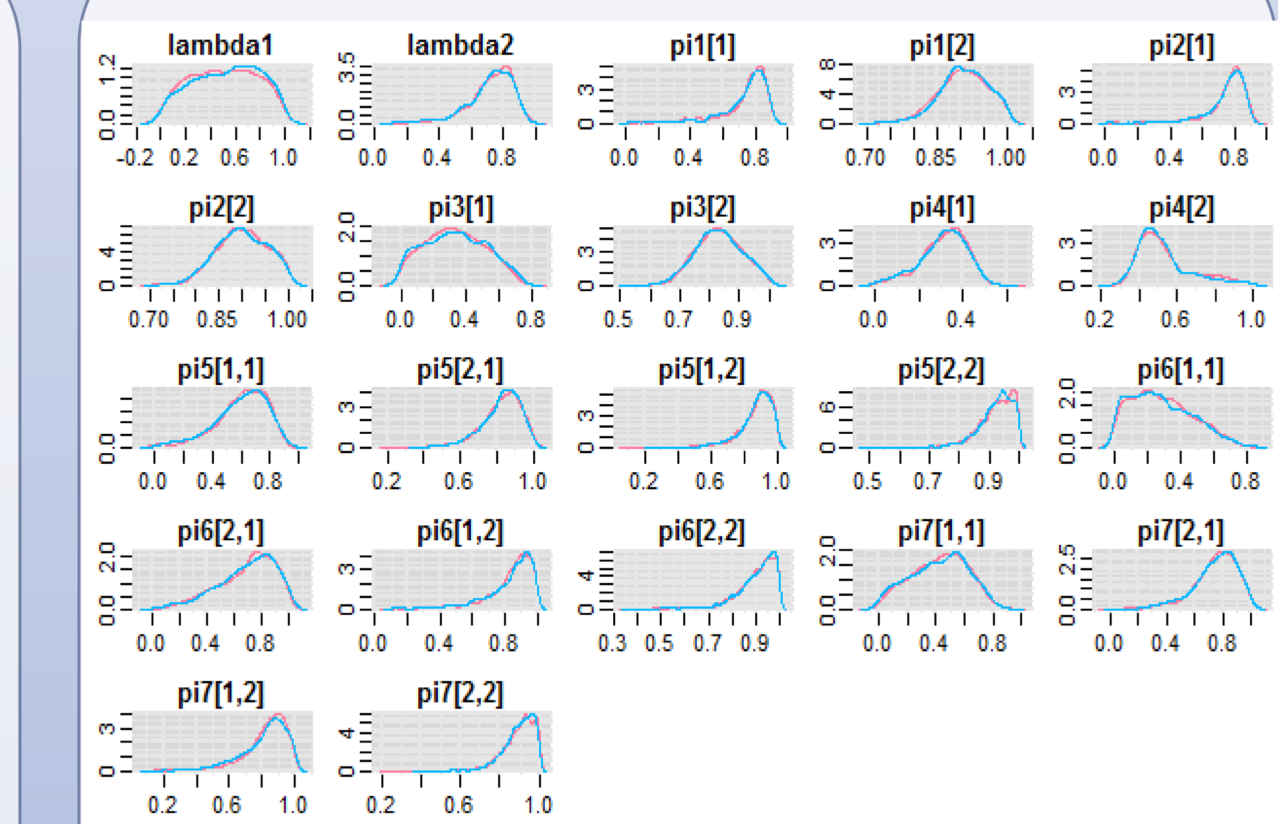
We tested the program using simulated data and generated trace plots. The convergence was good based on visualization and Potential Scale Reduction Factor (PSRF) criteria.

RESULTS & DISCUSSION

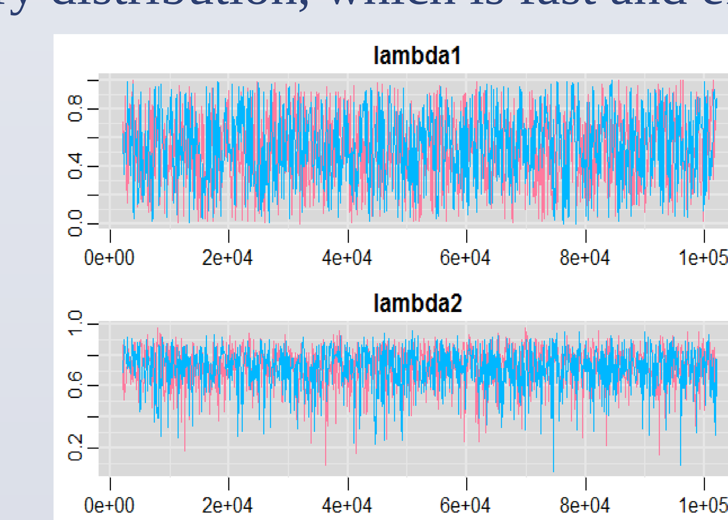




We applied our model on the real-world data from an introductory statistics educational App (SMILES). As shown above, we built an evidence model defining the relationship between competency and task variables based on the question designs about measure of central tendency (Mean & Median) from the App. In this case, ***θ1*** is about the their interpretations (e.g. median as 50th percentile & mean as numerical average). ***θ2*** is about the effect of outliers on them (e.g. mean is sensitive & median is robust). There is no attribute of individual in this case since the web logs did not record any such information. We ran the model on the preprocessed data and got posterior statistics from the perspectives of both conceptual competency and task response.



Again, the convergence is good based on visualization and PSRF criteria. We could see that two chains are tightly mixed with each other in all cases. Also, each density plot has a peak which consists with the result of descriptive statistics and maximum likelihood methods. Another highlight is that to fit our model appeared to require only a few thousand iterations of the MCMC to reach the stationary distribution, which is fast and efficient.



We noticed that the posterior distribution of lamda1 is not as informative as others. We reviewed its trace plot and found that the convergence was good. Thus, it is likely that tasks associated with student interpretation of the mean and median are less informative about underlying conceptual competency than tasks associated with the robustness of these measures.

ACKNOWLEDGEMENT & REFERENCE

We thank Dr. J Pearl who invented Bayesian Network and Dr. R Almond, Dr. R Levy who provided tutorial about Bayesian Networks in Educational Assessment (http://pluto.coe.fsu.edu/BNinEA/NCMETutorial/) that greatly assisted the research. We also thank SMILES (https://www.causeweb.org/smiles/team) project group for providing data. We thank JAGS for providing MCMC sampler.

CONTACT

Wentao Yan wentaoyan118@163.com