

Homework 1 - Perspectives on Computational Modeling - Winter'18

1. Classify a model from a journal.

a) Find a theoretical or statistical model from a recently published article (no earlier than 2013) in either the American Economic Review, American Journal of Political Science or the American Journal of Sociology.

I will be using a model from the American Journal of Sociology.

b) Give a detailed citation of the article.

Jæger, M. M., & Breen, R. (2016). A dynamic model of cultural reproduction. American Journal of Sociology, 121(4), 1079-1115.

c) Write down the mathematical or statistical model:

$$C_c = \beta_1 \theta_p + \beta_2 S_p + \beta_3 X_p + \beta_4 A_C + L_c$$

C - Child's cultural capital

β_1 - The return to parental investments in child's cultural capital

β_2 - The rate of transfer of cultural capital from parents to children

S_p - Parent's total stock of cultural capital

θ - The amount that the parents actively invest in their child

X_p - Parent's socioeconomic resources

A_c - The child's academic ability

L_c - Luck

This equation attempts to model how a child's cultural capital increments from parental contributions and interactions. Cultural capital is measured in arbitrary units, and is used to represent the number of cultural items that the child possesses at any given time.

d) Endogenous and exogenous variables:

Exogenous Variables:

- 1) Parent's total stock of cultural capital
- 2) The child's academic ability
- 3) Luck
- 4) The amount that the parents actively invest in their child
- 5) Parent's socioeconomic resources

Endogenous Variables:

- 1) Child's cultural capital.
- 2) The return to parental investments in child's cultural capital.
- 3) The rate of transfer of cultural capital from parents to children.

e) Classify the model:

- The model is stochastic. Usually, the same output will be generated if all the parameter values and initial conditions are fed. However, this model contains a random Luck element that will take different values during each iteration of the model.
- The model is dynamic. It tracks change over time as it models the intergenerational transfer of culture.
- The model is linear. There are no higher than 1 power terms in the equation.

f) List a variable or feature that you think the model is missing that might be valuable.

The presented model attempts to forecast how cultural capital is transmitted across generations. It accounts for both the receptiveness of the child (through the child's academic ability) as well as the several factors influencing the parent's propagation of cultural capital. However, cultural transmission across generations is presumably also influenced by cultural interactions between the child and their peers, or non-parent cultural interactions. The present model assumes that cultural transmission occurs primarily through parent-child interactions. This is not necessarily true: including an additional variable that accounts for cultural capital exchanged through interactions with friends and peers will be a valuable addition to the model. This variable can be created by analyzing both - the number of peer interactions occurring with culturally entrenched individuals as well as the nature of those interactions.

2. Build your own model

a) Write down a model of whether someone decides to get married:

Using logistic regression:

$$Y = \Pr(Y = 1 | X) = \beta_0 + k\beta_1 + q\beta_2 + p\beta_3 + o\beta_4 + w\beta_5 + n\beta_6 + z\beta_7 + \varepsilon$$

k - Level of education (categorical variable)

q - Ethnicity (categorical variable)

p - Age (continuous numerical variable)

o - Religiosity (treated as a continuous numerical variable. This will be in response to a question such as "On a scale of 1-10, how religious do you think you are?")

w - Income Level (categorical variable)

n - Cumulative net-worth (categorical variable)

z - Level of debt (continuous variable)

ε - Life circumstances/random variation in factors influencing marriage decision/luck (continuous variable)

b) At-least one of the dependent endogenous variables must be 1=get married or 0=not get married.

Y is a dependent endogenous variable that can take either a value of 1 or 0.

c) Make sure that your model is a complete data generating process.

This model can simulate data given all the parameters and relationships. All of this data can be obtained from participants using a survey administered to a random sampling of an American population cross-section. The random sampling will ensure that there is sufficient variation in the demographic and socioeconomic characteristics being used to model the decision to marry or not.

d) What do you think are the key factors that influence this outcome?

A range of demographic features, most importantly money-related characteristics of the person in question, determine the decision to marry or not. In my opinion, the decision to marry is most

significantly influenced by age and finances. In some instances, religiosity of the individual can galvanize the decision to marry, but in Western culture, religion does not solely drive the decision to marry as finances and age do. There will be systematic variation in ethnicity and decision to marry, as some cultures are more insistent on the idea than others. This ethnicity component is treated independently from the religious component for simplicity's sake. Lastly, I would expect education level to influence marriage decisions as these are ultimately linked to financial traits too.

e) Why did you decide on these factors and not others?

I attempted to find the most fundamental driving forces behind the decision to marry. The institution of marriage does serve a biological need of mating but at the same time, it has the robust social function of consolidating resources of two individuals. All cultures intertwine marriage and finances in numerous ways, and this led me to believe that at some fundamental level, the decision to marry cannot be triggered until the individual perceives their financial situation to be healthy, stable or in a desired state. Including financial variables subsumes other relevant variables that co-vary with finance: for instance, if one has a high net worth, a large annual salary, and a high level of education, then it is also likely that they grew up in a city, are attractive (attractive people have been shown to earn more than their counterparts), and are well-cultured. However, the reverse relation is not necessarily true (not all attractive people are rich), and hence I thought it made more sense to include those variables we know correlate with other marriage-influencing traits - such as socioeconomic variables. Additionally, the model would be too complicated if all relevant variables were included in the model. The present model attempts to be as simple as possible by capturing variation in those critical variables that are correlated with other, relevant variables but are more readily observable. Hence, both a need for simplicity and the need to be a complete data generating process are used to pick relevant variables.

f) How could you do a preliminary test of whether your factors are significant in real life?

A digitally-enhanced survey can be conducted on the internet to quickly gain people's ascribed importance to the host of variables included in the model here. A binary regression can be performed on the subsequent data to assess if any of the factors are trending towards significance. Once a preliminary test has succeeded, the follow-up full scale survey explicated above can be deployed to construct the model.