

Homework 01: Model Building and Model Selection/Fitting

MACSS 18

Di Tong

Building models

Deviant aggressive behavior

1. What social policy would be appropriate to reduce deviant aggressive behavior if Theory I were correct? Theory II? Theory III? Theory IV?

If theory I were correct, social policies that harshly penalize deviant aggressive behaviors would be appropriate to reduce deviant aggressive behaviors. In this case individuals would learn from experience that they might end up in death penalty if they shoot someone, and hence avoid such deviant aggressive behavior.

If theory II gained credibility, social policies that could lead to the lessening of deviant aggressive behaviors would be improving the supportive systems and institutions that help alleviate personal frustrations and build more harmonious relationship between authority and non-authority figures in the same sphere. For instance, there could be a policy requiring company authorities to express care for their employees from time to times and requiring companies to provide employees with high quality free consultation services and wellness workshops that help with problems from minor personal frustrations to serious mental health challenges. Under such policy, the employees' potential desire to express hostility to their bosses due to their personal frustrations can be dissipated. As a result, there would be less deviant aggressive behaviors.

If theory III had the strongest explanatory power, affirmative social policies that aim at systematically eliminate social discrimination would reduce deviant aggressive behavior. Suppose there's a group of people suffer from racial discrimination in terms of education and employment opportunity and therefore get stuck in poverty and tend to act deviantly and

aggressively. If a policy ensures their equal rights in receiving education and getting employment, they would start to profit from conforming to social rules. Then their rational action would switch from social deviant behaviors to rule-abiding behaviors.

Regarding theory IV, effective social policies to diminish deviant aggressive behaviors would be controlling the influence of deviant subcultures. Specifically, it could be done through limiting access to websites that propagate deviant subcultures and illegalize social organizations that promote deviant cultures and actions. Without contact with deviant subcultures, individuals won't have the opportunity to be socialized into the deviant role. Therefore, deviant aggressive behaviors would decrease.

2. During the past ten years, American society has been running a series of “experiments” with deviant aggressive behavior. Take any one of these experiments (e.g. #MeToo, mass shootings, political rhetoric) and discuss what we have learned about the four theories from this series of experiments.

The Stoneman Douglas high school shooting is a case of “experiment” in terms of deviant aggressive behavior. It is one of the deadliest shooting at high school in US. The shooting happened on February 14, 2018 in the Marjory Stoneman Douglas High School in Parkland, Florida. 17 students and staff members were killed and seventeen others were injured. Nikolas Cruz, a nineteen-year-old former student of this school was the shooter.

Except for the first theory, all theories listed here can account for Nikolas Cruz's deviant aggressive behavior to some extent. Among them theory II seems to be the most plausible one. Nikolas Cruz is an orphan and his adoptive parents died in his childhood and teenage years. He has suffered from and struggled with a range of mental health challenges, including depression, autism, and attention deficit hyperactivity disorder (ADHD). He has been expelled from schools for disciplinary reasons several times, including the Marjory Stoneman Douglas High School where the bloody shooting happened. Though desiring friends and care just like other teenagers,

Nikolas Cruz has always been ignored and marginalized. Clearly he fits the condition of being very frustrated in personal life according to theory II. He once told his friend that he was badly treated in high school and “going to go there and shoot it up” some day. For a student, the school can be viewed as personal authority figures. A frustrated student would tend to be angry about a school that once ignored, denied and expelled him. Hence theory II can plausibly explain Nikolas Cruz’s behavior: it is a symbolic expression of hostility toward personal authority figures.

The evidence for theory II also supports theory III to some degree. Because of his mental problems, Nikolas Cruz has always been marginalized by teachers and classmates and even expelled by schools. Such experience strengthened his erratic tendency and hence also strengthened his bad reputation, resulting in endless marginalization and expulsion of him. Therefore, just as theory III stated, as an oppressed individual, it is a rational action for Nikolas Cruz to not conform to social rules.

In addition, theory IV could also contribute to explain Nikolas Cruz’s behavior in that he is quit often in contact with violence-related cultures. He is a great fan of guns and armies and enjoyed hunting. He owns gun magazines with swastikas carved in them. He was a member of the Junior Reserve Officers' Training Corps and his school's varsity air rifle team. At age 18, he purchased an AR-15 style semi-automatic rifle for himself and had aspirations to join the military. Hence, we may infer that Nikolas Cruz has been socialized into the role of deviant aggressive behavior through exposure to violence-related cultures.

Theory I, however, does not work in this case. Apparently Nikolas Cruz had been punished for his erratic behavior for years under different circumstances, such as being banned from entering classrooms and expelled from school. Besides, there’s no evidence showing that he had ever received rewards for his deviant actions. Yet he not only did not learn from his experience to avoid deviant behavior, but also went on further in his deviant tendency to the extreme of mass shooting.

Waiting until the last minute

(a) Ask yourself why the observation might be true and write down your explanations.

To start with, people tend to seek instant pleasure and therefore avoid doing things they dislike if they could. As a result, they would not start working on a task they don't enjoy until they have to start. Another possible explanation stems from efficiency concern. Due to the pressure of possibly getting into trouble for not finishing work on time, people tend to be more efficient at the last minute before the deadline. Hence, it would take less time for one to accomplish a task if he/she starts at the last minute. Out of the purpose of efficiency, people would choose to wait until the last minute. In addition, under some circumstances, people might be waiting for inspiration to do their work better. As inspiration often comes in an emergency, that is, at the last minute before deadline. In the end people who are inspiration seekers appear to start late. Besides, people might learn from their experience that accomplishing task at the last minute would bring them extra pleasure. To finish a task in a rather short period right before the deadline is very risky and challenging. Such attributes make this process an exciting game that looks appealing to risk lovers. They would enjoy extra excitement when succeed in tackling the challenge. They would also feel that they are extremely competent and cool to accomplish such a challenging task. Then when the next task comes, to regain all these extra pleasures they experienced in waiting until the last minute to work on a task, they would again, choose to procrastinate.

(b) Generalize the explanatory model – that is, induce the most general, abstract model you can produce that still has the original observation as a consequence.

This model builds on the explanation regarding efficiency. It serves to examine whether people wait until the last minute to do things is due to that they value efficiency and want to get their work done more efficiently in this way.

The outcome variable measures how early people start working on a task. Specifically, I would treat this earliness variable as a continuous variable resulting from the task deadline time subtracting people's starting time to work on the task:

$$\text{Earliness} = \text{deadline time} - \text{starting time}$$

The lower the value of earliness, the later people start working on a task. People who wait until the last minute to do things tend to have lower value in this outcome variable.

The predictor variable measures people's preference regarding efficiency. I would use a scale from 1-5 based on answers to a survey question: to what extent do you think efficiency is important in accomplishing a task? From 1 to 5 there would be: very unimportant, unimportant, not important and not unimportant as well, quite important and very important.

The model would also control for the basic demographic characteristics such as age, gender, income, occupation, etc. as well as potential influencing factors stemmed from other hypotheses/theories.

(c) Induce an alternative model that also has the original observation as a consequence.

This model is based on the explanation regarding instant pleasure. It serves to examine whether people procrastinate until the last minute to do things is owing to that they do not enjoy doing them.

The outcome variable measures how early people start working on a task. Specifically, I would treat this earliness variable as a continuous variable resulting from the task deadline time subtracting people's starting time to work on the task:

$$\text{Earliness} = \text{deadline time} - \text{starting time}$$

The lower the value of earliness, the later people start working on a task. People who wait until the last minute to do things tend to have lower value in this outcome variable.

The predictor variable measures the degree to which people enjoy working on the task concerned in the outcome variable. Practically it would be a scale from 1 to 5, corresponding to not enjoy at all, not really enjoy, don't have a clear inclination, slightly enjoy and enjoy very much.

The model would also control for the basic demographic characteristics such as age, gender, income, occupation, etc. as well as potential influencing factors stemmed from other hypotheses/theories.

(d) For each of the two general models produced in (b) and (c), derive two interesting predictions (four predictions in total). Be sure the logical connection between your model and your predictions is explicitly stated and that any assumed facts concerning the world are made explicit.

For the general model in (b), one prediction could be that people who value efficiency more tend to wait until the last minute to do things. If, under the model in (b), the outcome variable is negatively correlated with the predictor, we gain evidence for a connection between people's preference over efficiency and their procrastination behavior. Another prediction could be that whether people value efficiency or not would not affect their procrastination behavior. If we do not find any statistically significant correlation between the outcome variable and the predictor under the model in (b), we could make the prediction that people's preference over efficiency is not a plausible explanation to account for their procrastination behavior.

For the general model in (c), one prediction could be that people who enjoy less working on certain task would start on the task later. If, based on the model in (c), data show that the outcome variable earliness is positively correlated with the predictor, it would count as evidence for the claim that people who enjoy a task to a lower degree tend to start on the task later. Another counter prediction could be that whether people enjoy working on certain task or not does not affect their procrastination tendency on the task. If data do not show any statistically

significant correlation between the predictor and the outcome variable based on the model in (c), the hypothesis regarding instant pleasure might not have explanatory power here.

Selecting and fitting a model

1. For each part, indicate whether we would generally expect the performance of a flexible statistical learning method to be better or worse than an inflexible method. Justify your answer.
 - a. A flexible method will generally perform better than an inflexible method, because an extremely large sample size enables flexible models to fit the data closer and at the same time produce accurate estimate.
 - b. A flexible method will generally perform worse than an inflexible method, because a flexible method needs a very large number of observations to obtain an accurate estimate of f and a small number of observations will bring about overfitting problem.
 - c. A flexible method will generally perform better than an inflexible method better, because more flexible methods can generate a much wider range of possible shapes, especially the non-linear shapes, to estimate f , while the inflexible method is more suitable for more linear relationships.
 - d. A flexible method will generally perform worse than an inflexible method, because it will fit to the noise and hence create higher variance.

See the solution to Problem 2 and 3 in the Jupyter Notebook file.