Perspective on Computational Modeling
Homework 1: Model Building and Model Selection/Fitting
Tianyue Niu

Question 1: Deviant Aggressive Behavior (word count: 797)

Theory I is similar to B.F skinner's theory of operant conditioning in psychology. The theory states that reinforcements and punishments can act to strengthen, or weaken a behavior during learning. For example, if people receive positive reinforcements (such as recognition, smile, or positive feedbacks) for being kind to others, then they will stay kind to others. On the other hand, if they receive positive reinforcement for stealing (for example, obtaining a brand new iPhone at zero cost) without punishment, then they will steal in the future. Given this theory, deviant aggressive behaviors should be reduced through rewarding pro-social behaviors, and punishing anti-social deviant behaviors. To encourage pro-social behaviors nationally, policy such as increasing welfare for citizens with no criminal records can be implemented. For example, hypothetically, a state can award citizens with no-criminal record with higher tax-refund to reinforce lawful behaviors. On the other hand, punishing deviant behaviors can also lead to reduced crime. For example, a state could potentially increase sentence terms for all kinds of crime or legalize death penalty to make the costs of conducting a crime higher. Policy such as expanding the police department can also be effective in that it increases the chance of punishment for unlawful behaviors.

Theory II states that the source of deviant behaviors is frustration with one's own life. If this were true, keeping everyone in the society satisfied would effectively reduce deviant behaviors. To accomplish this, a government can attempt to increase social welfare through several different angles such as unemployment benefit, healthcare, education, and tax. First, policies such as increasing unemployment benefits in the form of more compensation per payment period or longer benefits duration can help increase unemployed people's general satisfaction with life. Second, universal healthcare can also be beneficial in that people would worry less about falling sick and not being able to pay for their hospital bills. Thirdly, adults with children would likely become more satisfied with life if their children are receiving good education. Therefore, subsidizing public schools for better teaching staff and better infrastructure can also lead to increased satisfaction among citizens. Lastly, decreased consumption tax and payroll tax can also make citizens more satisfied because they would have more buying power to purchase better food, better clothes, and better services for themselves and their families. These

policies all contribute to increase people's satisfaction with life. According to theory II, this should effectively reduce frustration and deviant aggressive behaviors.

Theory III states that oppressed individuals that are systematically discriminated by the society are more likely to become deviant. If this theory were correct, increasing equality and decreasing discrimination should reduce deviant behaviors. Policies can be implemented at different levels, or in different sectors, to promote equality. For example, in terms of education, admission to higher education should not discriminate applicants based on race, gender, or country of origin. This can be reinforced by implementing policies such as not asking for applicant's personal information during application, or requiring schools to block applicants' personal information before making an admission decision. The same policy can be implemented in the job market as well. Job candidates should not be discriminated based on race or gender. Policy such as banning employers from asking candidates to 'self-identify' their demographic information 'voluntarily' can might to increased equality. Nationally speaking, promoting equality in terms of income distribution can be achieved through policy such as implementing a progressive tax system, in which one's payroll tax is proportional to one's income. These policies help promote equality and reduce oppression against certain groups of citizens. According to Theory III, these policies should contribute to reduce crime.

Lastly, Theory IV states that individuals can become deviant after exposure to a deviant culture. If this were the main cause of deviant behaviors, it would be challenging to correct delinquency because it is hard to control people's contact with their surrounding culture. One policy that can be implemented is to prevent children from watching violent movies and TV shows or playing violent games. However, it can be very difficult to implement due to the prevalence of media. Another way to reduce people's, or at least children's exposure to a deviant culture is to establish a better foster care system. For example, for children born into violent families, if the children were constantly abused, then actions should be taken to either convince their caretakers to change their behaviors or transfer these children into better hands. However, since leaving one's biological parents can be psychological traumatizing, policies like this need to be implemented with extreme cautions. If the foster care system is well-established, then this could potentially decrease a certain group of children's exposure to violent culture. According to Theory IV, this could lead to less deviant aggressive behaviors.

Question 2: Procrastination

a. Why do people procrastinate?

People may procrastinate until the last minute for a variety of reasons. First, they might lack the confidence to do well in the task, so they avoid doing it. Second, they might not enjoy doing the task, so they wait until they have to do it to start the task. Third, some people might lack the ability to focus for a long time, so they would procrastinate until the last minute because otherwise, they wouldn't have enough motivation to focus on the task without diverging.

b. Generalize the explanatory model

Challenging tasks are often less enjoyable and more likely to incur negative emotions, so people are more likely to procrastinate challenging tasks than easier tasks.

c. Introduce another model

People's ability to focus negatively correlates with procrastination time.

d. For each model, make two interesting predictions.

For model 1, the first prediction is that, suppose a group of people are each given an easy task (list 3 of your favorite colors) and a harder task (list the assumptions of linear regression) to do with the same deadline, the majority of the group would choose to finish the easier task first.

The second prediction is that, suppose we have a group of people divided into treatment and control group, and we ask them to finish the same math problem set given a deadline (e.g. two days later). For the control group, no instruction will be given. For the treatment group, an instructor would walk them through every question and let them know the best way to solve them. My prediction according to model 1 is that the treatment group would hand-in the problem set earlier than the control group.

For model 2, the first prediction is that, when we give children with ADHD and children without ADHD the exact same math problem set with the same deadline, on average, the children without ADHD is going to hand in the problem set earlier than the children with ADHD.

The second prediction is that, suppose we give a task to two groups of people with the same deadline, the first group of people live alone, and the second group of people live with their best friends. Assuming that those living with their best friends are constantly harassed by their friends to hang out with them, my prediction is that, because the second group are less able to focus on the task, on average, they will turn in the task later than the first group.

Question 3: Selecting and fitting a model

1) Flexible/Inflexible Method

a) *Sample size n is extremely large, and the number of predictors p is small.*

Since sample size is extremely large, the model would be less affected by outliers. In this case, we can use a flexible model to fit the data.

b) *Number of predictors p is extremely large, number of observations n is small.*

An inflexible method should be used because number of observations n is small, so using a flexible model would likely overfit.
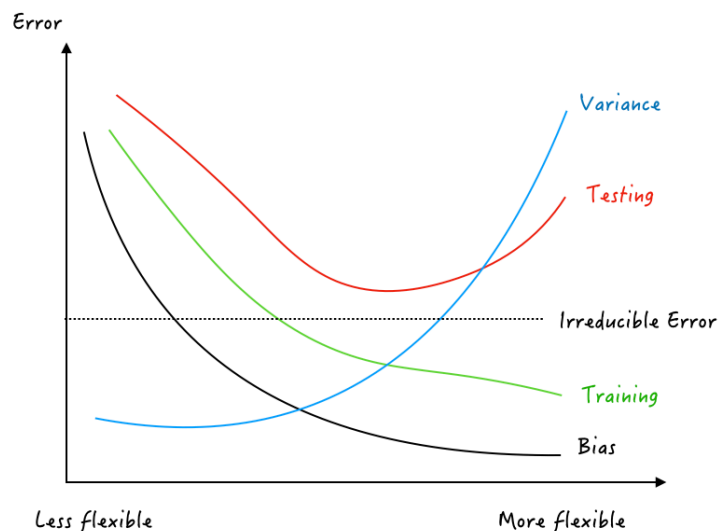
c) *Relationship between the predictors and response is highly non-linear.*

We should use flexible method because the relationship is highly non-linear, and so we need a flexible model to try its best to fit the non-linear data.

d) *The variance of the error terms $\sigma^2 = Var(\varepsilon)$ is extremely high*

We should use inflexible method because otherwise, overfitting would occur (as the model would try to fit the high variance/noise).

2) Explain why each of the five curves has the shape it has.

Variance: as the model complexity increases, variance would become higher because the model would try to fit every data point, which means outliers would also be fitted, therefore leading to higher variance.

Training: error for training would keep decreasing as the model becomes more complex because, as m mentioned above, the model is going to try to fit every data point, so the error made would decrease.

Testing: error for testing sample would first decrease and then increase. It would first decrease because the model would learn more about the data as the model becomes more complex, and so it would become better at predicting/generalizing. However, as the model becomes overly complex, overfitting occurs and the model's to generalize to other data set (other than the sample) decreases, and so error increases.

Bias: bias would monotonically decrease because as the model becomes more complex, the estimation (for the training sample) will become more accurate/closer to the mean.

Irreducible error: true variance of the data is irreducible and it should be constant despite the model complexity.