Problem Set 1

Pete Cuppernull

1/19/2020

1) Deviant aggressive behavior - What social policy might be appropriate to reduce deviant aggressive behavior if Theory I were correct? Theory II? Theory III? Theory IV?

Considering the four provided theories, different social policies would be appropriate to combat deviant aggressive behavior. In Theory I, where deviant aggressive behavior is learned from experience, I propose that resources should be allocated towards improving the quality of primary and secondary education, with an emphasis on smaller class sizes. Considering that schools are a formative environment for developing social skills, this could be a critical entry point for punishing deviant behavior and rewarding acceptable behavior. I emphasize smaller classes because it seems that much of this feedback would have to take place at the interpersonal level (for example, with increased interactions between teachers and students). It is important to note that because of this, other reforms to schools might not have an effect on deviant aggressive behavior – for example, providing each student with their own computer would likely not be effective because it would have a limited impact on punishing and rewarding certain behaviors.

For Theory II, which describes deviant aggressive behavior as "a symbolic expression of hostility toward personal authority figures", it would be necessary to first allocate resources towards a public survey to better understand personal sources of frustration – per the theory, these sources of frustration lead to the deviant aggressive behavior. Once these sources are identified, I would direct public resources towards addressing them. As an example, if wage insecurity was discovered to be a source of frustration, a minimum wage increase could potentially help alleviate deviant aggressive behavior.

In the case of Theory III, where deviant aggressive behavior is "the rational action of oppressed individuals", social policies that limit the degree to which people feel oppressed would be appropriate. For example, the implementation of more frequent and robust community town hall events could improve the degree to which individual voices are heard in local politics. If individuals feel less oppressed once given a stronger voice in politics, then according to the theory, deviant aggressive behavior should decrease. Similar reforms that bolster the voice on the individual in local politics could also be useful – for example, local governments could provide extra transportation from low income neighborhoods to polling stations on election day.

Finally, in Theory IV, where deviant aggressive behavior is "a social role" and "individuals are socialized into the role through contact with a deviant subculture", it would be appropriate to implement policies which allow for the productive expression of deviant behavior and the deviant subculture. One example of such a reform would be to implement graffiti walls in public parks where graffiti artists can legally paint, effectively celebrating grafitti rather than punishing it (a policy along these lines was implemented in Venice Beach, California, and the graffiti walls are now a popular tourist attraction). Beyond this, local governments could sponsor youth sports tournaments to allow individuals to compete in a productive environment.

2) Waiting until the last minute

A. Ask yourself why the observation might be true and write down your explanations.

People might wait to do things until the last minute because they prioritize tasks with earlier deadlines. Things to be accomplished by a later date are deferred until more immediate tasks are completed. Alternatively, it is possible that individuals tend to prioritize tasks which they find most enjoyable. Thus, less enjoyable tasks are left until the last minute until they must be prioritized due to an impending deadline.

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.

People seek immediate gratification. Work that can be delayed will be delayed, and of the work that must be accomplished, the most enjoyable work is prioritized.

C. Induce an alternative model that also has the original observation as a consequence.

People believe that they produce higher quality work under pressure. Having a limited amount of time to complete a task makes it a higher-pressure task – so, individuals strategically wait to the last minute as a means of boosting their performance.

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.

- Model B predictions:
  1. People tend to save less money than they otherwise could. Here, I assume that spending money is more enjoyable than saving money. Money is spent as a means of acheiving immediate gratification.
  2. Individuals with dietary restrictions sometimes eat foods known to cause them harm. For exakple, someone who is lactose intolerant might occassionally eat pizza – the immediate gratification that comes from eating the pizza is prioritized over the later consequences of doing so.
- Model C predictions:
  1. People perform complicated tasks in front of others. I assume that the presence of observers raises the stakes of a given action – examples of this might include sporting events or theatrical performances.
  2. People take on prestigious tasks. Assuming that there is greater pressure involved with completing prestigious tasks, people will take on the extra work associated with these tasks.

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. The sample size n is extremely large, and the number of predictors p is small.

   Flexible. With a larger sample size, we reduce the problem of potentially overfitting a model. Thus, we can exploit the predictive power of a flexible model with a minimized concern of overfitting.

b. The number of predictors p is extremely large, and the number of observations n is small.

   Inflexible. Here, we would be very concerned with overfitting due to the small sample size and the high number of predictors. Therefore, we would prefer the restricted nature of an inflexible model that is less prone to overfitting.

c. The relationship between the predictors and response is highly non-linear.

   In this example, we would prefer a flexible model because an inflexible model would be biased in predicting the non-linear relationship. However, we would have to be conscious of not overfitting the model.

d. The variance of the error terms is extremely high.

   Inflexible. Flexible models would attempt to explain the high variance of the error terms as part of the model parameters. This would detract from the ability to model the true data generating process.

2. Bias-variance: Explain why each of the five curves has the shape it has.

- Bias – Bias decreases as models become more complex because the flexibility of more complex models allows the predicted values of Y to be closer to the actual values of Y. The logarithmic shape of the curve indicates that bias becomes increasingly difficult to reduce as models become more complex.
- Variance – Variance increases as models become more flexible because more of the noise in the training data is captured in the model. Thus, the model's predictions for test data are further from the true values of these test data.
- Training Error – Similar to bias, the training error will decrease logarithmically with more flexible models because these flexible models can fit more of the noise in the training data.
- Test Error – As models become more flexible, test error first decreases, then reaches a minimum and begins to increase. Very inflexible models will exhibit higher test error because they have not been sufficiently trained to model the data generating process. Test error will decrease to an extent as models because more flexible, but will begin to increase once these models start to overfit the training data. Overfitting causes an increase in test error.
- Irreducible Error – Irreducible error exists separate from a model and cannot be reduced via proper model design; thus, it is a constant that does not change as models become more or less flexible.