

Sruti Kanthan  
MACS 30100  
Jan 18, 2020  
Homework 1: Model Building and Model

## **Building models**

### **Deviant aggressive behavior**

Answer the following question in 500-800 words:

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

The four given theories describe potential sources of socially deviant aggressive behavior, and each point would require separate but interrelated policy initiatives to address them.

The first of these theories, Theory I, states that deviant aggression is a learned behavior reinforced through operant conditioning. In other words, learned behavior is a result of incentivizing aggression through rewards for aggressive behavior, and the individual learns to pair aggression with impending benefits. If this theory were correct, disincentivizing aggressive behavior is key. If aggressive behavior is met with punishment instead, then (through the lens of this theory), the behavior would stop. However, the nature of the punishment used is crucial; research shows that reacting to aggressive behavior with aggression, as a punitive measure, actually exacerbates an individual's aggression and can lead to mental and behavioral dysfunction later in life (Miller-Perrin & Perrin, 2018). If Theory I correctly gauges the source of aggressive behavior, social policies that better subsidize mental health care would help families find therapists who can help them learn how to address aggression, particularly between ages 0-5 when children are the most developmentally malleable and receptive to psychosocial interventions (Miller-Perrin & Perrin, 2018). Policies that institute harsher legal penalties for deviant aggression in adulthood--coupled with rehabilitative measures--would also help break the association of aggression with reward.

The second theory, Theory II, posits that deviant aggressive behavior is the result of displacement: the transfer of feelings from somebody onto a more socially acceptable (or accessible) target. For instance, from the perspective of this theory, an individual who is angry at a government official might go home and yell at his wife instead. In other words, the individual feels angry but unable to communicate this to the person in question, and instead takes out his anger on someone else. If this theory is correct, it might be difficult to generate a social policy to counter it because it is difficult to predict who is vulnerable to becoming the target of an aggressor's displaced rage. Some of the policies from the previous paragraph, such as making mental healthcare more accessible, could help people understand the true source of their anger and how to more healthily manage it. Other policies might encourage better employer-employee relations through empathy trainings, to help foster better relationships in the workplace and therefore reduce the risk of anger towards an employer.

The third theory, Theory III, says that aggression is the natural, expected reaction of social inequity – individuals who have been historically marginalized or oppressed are far more likely to subvert social rules given that these rules cause them harm. The most obvious way to address this, assuming that Theory III is correct, is to encourage social policies that right social inequities. If social rules change such that they actually benefit oppressed people, then people will be incentivized to follow those rules. For example, black individuals are disproportionately affected by lead exposure in Chicago, pointing to the lingering effects of residential segregation (Muller & Winter, 2018), which could generate understandable anger (or, “deviant aggression”) from residents. However, rallying for policies that require more stringent lead checks could help mitigate this aggression. Alternatively (and controversially), some social policies might encourage creating programs to help oppressed individuals learn how to protest peacefully, without exhibiting or resorting to deviant aggressive behavior.

The fourth and final theory, Theory IV, states that aggressive behavior is the result of socialization, after

contact with a “deviant subculture”. The same social policies that were explicated above apply here as well-mandating social policies that generate accessible programs to help prevent this behavior from forming, or buffer the effects of socialized aggression. Some might argue that censoring these channels is counterproductive and even induces a backfire effect; others feel that censorship is the only way to deplatform deviantly aggressive individuals online. Additionally, moderation on some of these forums (e.g. 4chan and reddit) or minimum age requirements for joining might help, but this is unlikely given that it would be easy to circumvent these conditions.

### **Waiting until the last minute**

People often do things at the last minute (students turning in papers, professors grading exams, and so on).

#### **a. Ask yourself why the observation might be true and write down your explanations.**

Procrastination could stem from a range of issues: “internal” extenuating circumstances, including a lack of motivation, diagnosed or undiagnosed mental health problems, perfectionism, etc. Some people might procrastinate because they enjoy the rush that comes with getting the assignment done at the last minute, potentially reflecting a self-regulation problem.

Procrastination could also result from “external” extenuating circumstances, such as physical illness, family obligations, or other logistical factors that take precedence over school.

Finally, procrastination could be viewed as the individual’s choice, despite these factors.

#### **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.**

We can parse our model such that certain personality traits (or, mental health issues) are more likely to be correlated with procrastination than others. In this setup, we would regress procrastination (e.g., scores on the GPS measure, Sirois et al., 2019) on our predictor variables, operationalized personality traits (e.g., dimensions from the Minnesota Multiphasic Personality Inventory and/or the NEO-PI).

$$\text{Procrastination}^{\wedge} = \beta_1(\text{Personality trait score}) + \beta_2(\text{Personality trait score}) \dots \beta_k X_k + \beta_0 + \epsilon$$

#### **c. Induce an alternative model that also has the original observation as a consequence.**

Alternative model: An individual’s locus of control (using the Rotter scale: Rotter, 1966) and age can better predict whether or not they will procrastinate.

$$\text{Procrastination}^{\wedge} = \beta_1(\text{locus of control score}) + \beta_2(\text{age}) + \beta_0 + \epsilon$$

#### **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.**

General model predictions:

1) Introverts are less likely to procrastinate than extroverts.

Plenty of research suggests that individuals who score high on the extravert dimension of the NEO-PI tend to be proportionally more sensation-seeking than their introverted counterparts (e.g., Aluja, Garcia, & Garcia, 2003). Therefore, it stands to reason that extraverts might be more likely to seek or experience the “thrill” associated with turning in an assignment at the nick of time.

2) Individuals who score high on the psychopathic deviate measure on the MMPI are more likely to procrastinate. I am not counting this as a prediction for my second model, given that the MMPI describes

A body of research suggests that antisocial personality disorder is marked by, among other features, propensity for risk-taking, impulsivity, and sensation-seeking behavior (e.g., Johnson, 2019). Given that procrastination might result in part from difficulty self-regulating and seeking an adrenaline rush, people who fit the APD diagnosis might be more likely to procrastinate.

#### Alternative model predictions:

1) Younger individuals are more likely to procrastinate.

Research suggests that younger people have lower self-control and conscientiousness than their older counterparts, given that their prefrontal cortex is still developing; this could signal a higher propensity for procrastinating (Steel & Ferrari, 2013).

2) Individuals with external loci of control are less likely to procrastinate.

Individuals with an external locus of control are more likely to believe in fate, God, or other external determinants dictating their life outcomes, so it makes sense that these individuals might be less averse to procrastination because they feel that the outcome is out of their hands, either way. Individuals with an internal locus of control, on the other hand, might be more likely to approach tasks head-on because they feel a deeper sense of personal autonomy, responsibility, and control over their own actions.

Of course, this is situationally dependent and there are dozens of other factors that might explain why an individual would procrastinate.

#### **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.**

A flexible method would perform better than an inflexible model in this case, given that a large sample size helps buffer against the over-fitting risk that usually comes with using a flexible method.

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

An inflexible model would be better, because it is less likely to pick up patterns that turn out to be noise, lowering the risk of over-fitting the data (given the small amount of observations) and preventing generalizability.

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

A flexible model would be more adept at picking up on a non-linear relationship; inflexible models are better suited for linear relationships.

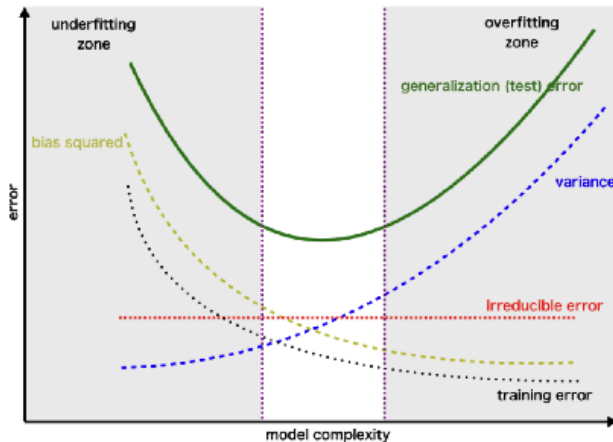
**d. The variance of the error terms is extremely high.**

High variance would mean a lot of noise, which a flexible method would be more likely to mistake for a meaningful pattern in the data and over-fit. It would be better to use an inflexible method in this case.

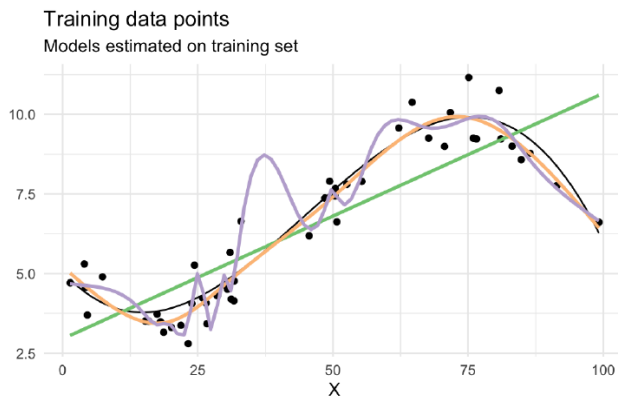
**2. Bias-variance:** Think about the graph including bias, variance, training error, test error, and irreducible error curves, moving from less flexible statistical learning methods towards more flexible approaches.

**a. Explain why each of the five curves has the shape it has.**

The following explanations are based on this figure:



- (1) **Bias:** As flexibility increases, bias decreases. This means that when the model is inflexible, bias is larger, decreasing as the model becomes more flexible. In other words, as the model becomes more flexible, it starts to represent the complexity of the data more closely, reducing bias.
- (2) **Variance:** As flexibility increases, variance increases, indicating a bias-variance tradeoff. Here, as the model becomes less rigid, it overfits, over-weights noise, and loses generalizability. Variance can also be thought of as a gauge of overfit risk, and it makes sense that as the model increases in flexibility the risk of overfitting also increases.
- (3) **Training error:** As flexibility increases and the model better fits the training data set, training error decreases, following a pattern similar to bias'.
- (4): **Test error:** As flexibility increases, test error initially decreases and then increases (parabolic) because bias reduces more quickly than the variance increases (in other words, represents the bias-variance tradeoff). When the model learns from the training dataset, test error initially reduces, but then increases when the model begins to overfit the data.
- (5): **Irreducible error:** This stays constant as the model increases in flexibility, because it is generated by the dataset and would not be influenced by the model's flexibility.



\*If this prompt corresponds to the graph from section 2.6 of our class notes on model fitting\*:

- The data is not quite linear, so the linear regression model (green line) is inflexible, underfitting the data (low variance and high bias).
- The model corresponding to the purple line is too flexible, and overfits the data (low bias and high variance).
- The models corresponding to the orange and black lines more closely fit the data (lower bias, lower variance).

## References

- Aluja, A., Garcia, O., & Garcia, L. F. (2003). Relationships among extraversion, openness to experience, and sensation seeking. *Personality and Individual Differences*, 35(3), 671-680.
- Johnson, S. A. (2019). Understanding the violent personality: antisocial personality disorder, psychopathy, & sociopathy explored. *Forensic Res Criminol Int J*, 7(2), 76-88.
- Miller-Perrin, C., & Perrin, R. (2018). Physical punishment of children by US parents: moving beyond debate to promote children's health and well-being. *Psicologia: Reflexão e Crítica*, 31(1), 16.
- Muller, C., Sampson, R. J., & Winter, A. S. (2018). Environmental inequality: The social causes and consequences 263-282.
- Sirois, F. M., Yang, S., & van Eerde, W. (2019). Development and validation of the General Procrastination Scale (GPS-9): A short and reliable measure of trait procrastination. *Personality and Individual Differences*, 146, 26-33.
- Steel, P., & Ferrari, J. (2013). Sex, Education and Procrastination: An Epidemiological Study of Procrastinators' Characteristics from a Global Sample. *European Journal of Personality*, 27(1), 51-58.  
<http://dx.doi.org/10.1002/per.1851>
- Zarzycka, B., Liszewski, T., & Marzel, M. (2019). Religion and behavioral procrastination: Mediating effects of locus of control and content of prayer. *Current Psychology*, 1-10.