**Peer-graded Assignment: Ethical Issues in Statistics and Data Science (Fair ML Intro)**

How are the models that we’ve used in this course so far (and in this specialization more broadly) based on the process of induction?

Generalized Linear Models (GLMs) and nonparametric regression rely on the principle of induction, aiming to extract patterns from specific data examples to make accurate predictions on new, unseen cases. As highlighted in the introduction, induction is about "formulating general rules based on specific instances—rules that not only explain past cases but also extend to future, unseen scenarios."

GLMs and nonparametric regression focus on identifying the relationships between variables using training data, striving to capture patterns that can be generalized rather than merely memorizing the data. These models then apply the derived patterns to predict outcomes for new data points. This transition from specific examples to general rules, which can be applied to future cases, is at the heart of the inductive approach in machine learning.

“The fact that machine learning is ‘evidence-based” by no means ensures that it will lead to accurate, reliable, or fair decisions.” Provide 2-3 examples of this claim.

a) **Historical Bias in Training Data**: The introduction points out that "Our historical examples of the relevant outcomes will almost always reflect historical prejudices against certain social groups, prevailing cultural stereotypes, and existing demographic inequalities." For instance, a hiring algorithm trained on past data might continue to replicate biases against women or minorities in specific industries.

b) **Feedback Loops**: The text explains how predictive policing algorithms can create self-fulfilling prophecies. If an algorithm predicts higher crime rates in particular areas, leading to increased policing, it may result in more arrests in those locations. This, in turn, reinforces the algorithm's predictions in subsequent iterations.

c) **Proxy Variables**: Even when protected attributes like race are excluded, other variables might serve as stand-ins. The text provides an example where the age at which someone starts programming correlates with gender. Therefore, using this seemingly neutral variable could still introduce gender bias into hiring decisions for programming roles.

Do you believe that Amazon's same-day delivery system mention on page 3-4 is unfair or unjust? Why or why not?

While there are valid points on both sides, I believe this system raises significant fairness concerns. Although Amazon argues that the system is driven by efficiency and cost rather than race, the pronounced racial disparities in service availability are alarming. The text highlights that in many cities, white residents were more than twice as likely as Black residents to live in neighborhoods that qualified for the service.

This situation reinforces existing inequalities, as predominantly white areas receive better service. Even if the system isn't deliberately discriminatory, its impact exacerbates systemic racial disparities. Given the long history of racial discrimination and segregation in the US, any system that results in noticeably different levels of service along racial lines is deeply problematic, regardless of whether race is an explicit factor.

What is the machine learning loop? Do you think that the machine learning loop, given in figure 1, applies to statistical modeling? Justify your answer.

I agree that the machine learning loop illustrated in Figure 1 also largely applies to statistical modeling. Both machine learning and traditional statistical modeling share the following steps:

1. **Measurement**: Gathering and preparing data
2. **Learning**: Fitting a model to the data
3. **Action**: Using the model for predictions or decision-making
4. **Feedback**: Evaluating the model's performance and making potential updates

The main distinction is that machine learning often emphasizes the iterative nature of this process, with models frequently updated based on new data and feedback. In contrast, traditional statistical modeling might involve less frequent updates. Despite this, the overall process is quite similar.

Both approaches encounter similar challenges, such as ensuring the data accurately reflects reality, ensuring the model generalizes effectively, and that the decisions or actions based on the model are both appropriate and fair. Additionally, the potential for feedback loops influencing future data and predictions is a concern for both machine learning and statistical modeling.