

Adjustment, is the idea of putting regressors into a linear model to investigate the role of a third variable on the relationship between another two. Since it is often the case that a third variable can distort, or confound if you will, the relationship between two others.

As an example, consider looking at lung cancer rates and breath mint usage. For the sake of completeness, imagine if you were looking at forced expiratory volume (a measure of lung function) and breath mint usage. If you found a statistically significant regression relationship, it wouldn't be wise to rush off to the newspapers with the headline "Breath mint usage causes shortness of breath!", for a variety of reasons. First off, even if the association is sound, you don't know that it's causal. But, more importantly in this case, the likely culprit is smoking habits. Smoking rates are likely related to both breath mint usage rates and lung function. How would you defend your finding against the accusation that it's just variability in smoking habits?

If your finding held up among non-smokers and smokers analyzed separately, then you might have something. In other words, people wouldn't even begin to believe this finding unless it held up while holding smoking status constant. That is the idea of adding a regression variable into a model as adjustment. The coefficient of interest is interpreted as the effect of the predictor on the response, holding the adjustment variable constant.

In this lecture, we'll use simulation to investigate how adding a regressor into a model addresses the idea of adjustment.