**BIOST 2049**

**Spring 2023**

**Research Article Dissection 2A – Assessing Regression Assumptions**

**Due 2/8/23 by 11:59 via the Assignment Link posted to Canvas**

Please read the article “Regression assumptions in clinical psychology research practice – a systematic review of common misconceptions” by Ernst and Albers found on Canvas and answer the questions below.

This is an individual assignment. You can discuss the article with your classmates, but the answers you submit should be your own.

Please be brief but write in complete sentences with appropriate interpretations (a yes or no answer will not get full credit). A few sentences is all that is necessary to answer each question.

1. What are the problematic situations that may arise when the regression assumptions are violated?

Desirable properties of regression estimators when assumptions are met are that estimators are expected to be near the true population parameters and are expected to consistently be near the true population parameters with minimal variance. When these assumptions are not met, we can’t guarantee these desirable properties of our estimates.

1. Is there anything different about the regression assumptions discussed in this paper compared to what we discussed in class? If so, state the differences.

There are no differences. The paper discusses the four OLS-assumptions we discussed in class: linearity, normality, homoscedasticity, and independence.

1. What are some common misconceptions about the assessment of the regression assumptions?

A major misconception is that covariates and outcomes have to be normally distributed. Another misconception is that features must be linearly related to the outcome, when in fact what makes a linear model a linear model is that the *coefficients* must be linearly related to the outcome.

1. List the best practices for assessing regression assumptions.

Graphical methods are preferred over hypothesis tests of assumptions. Residual plots are useful for homoscedasticity and linearity. QQ-plots are useful for checking normality, and autocorrelation of the residuals is useful for the independence assumption.

1. Were you surprised by the results of systematic review? Explain.

Yes. That 92% of research papers were observed to have unclear assumption checking means there is a good chance that some proportion of these papers have models with violated assumptions. Therefore, the results and conclusions contributed to the literature by these studies are less than ideal.