Quiz 5

1 – Which of the following is not a reason that multivariable modeling is preferred over single-variable modeling?

* Collinearity (automatically dealt with)
* Swamping (automatically dealt with)
* **Homoscedasticity**
* Interactions (can be included!)
* Pseudoreplication (can be accounted for)
* Elegance (yes)

**Explanation: I mentioned 6 reasons why multi-variable modeling is preferred: collinearity (dealt with), swamping (dealt with), interactions (can be included), pseudoreplication (can be accounted for), and elegance (it’s nicer). Homosecdasticity / heteroscedasticity can be dealt with in statistics, but this can be done in single variable models.**

2 – If collinearity exists among independent variables (X) that all influence the response variable (Y), what problem occurs when running analyses using only single X-variables?

* Swamping
* **Bias in estimates**
* Variance inflation
* Interactions

**Explanation: The key is that they influence your results! They are confounding variables. And if confounding variable are left out, they influence your results, and bias your estimates.**

3 – Multi-variable analyses overcome the problem mentioned in Question #2, but it introduces a new problem. What problem exists in multi-variable analyses when X-variables are collinear?

* Bias in estimates
* **Variance inflation**
* Interactions
* Swamping

**Explanation: And variance inflation happens when there is collinearity whether the variables are confounding or redundant.**

4 – Which of the following is not a symptom of collinearity among your X-variables?

* Model p-value is significant, but individual variables aren’t significant
* High r^2 among X-variables
* **Betas close to zero**
* High VIF scores
* Coefficients change when adding or removing variables from models’

**Explanation: betas close to zero can happen whether you have high collinearity or not. Otherwise, high collinearity can decrease p-values of individual variables, shows correlation between pairs of x-variables, have high VIF scores, or your coefficients change when variables are added or removed (due to bias from collinearity not being accounted for).**

5 – What is swamping?

* When the effect of one X-variable depends on the value of another X-variable
* **When the ability to detect the effect of one variable is masked by the effect of another variable**
* When two or more X-variables are correlated with each other
* When leaving one or more X-variables out of the model introduced bias in the estimated effects of other X-variables in the model

**Swamping – the ability to detect the effect of one variable is masked by the effect of another variable. By building multi-variable models, we can handle this potential problem.**