Question 1

Describe the difference between a moderating and a mediating variable. Which type of variable is more closely related to an interaction effect in regression or ANOVA? Why?

Solution:
If the third variable Z is intermediate in a causal sequence such that X causes Z and Z causes Y, then Z is a mediating variable; it is in a causal sequence X Z Y. But in moderation, there is not an indirect relation of X to Y but there is an interaction XZ that corresponds to a potentially different X to Y relation at values of Z. In analysis of variance (ANOVA) and regression, moderator effect can be represented as an interaction between a focal independent variable and a factor that specifies the appropriate conditions for its operation.

Question 2

Briefly (in one or two sentences) explain the idea of moderated mediation
Solution:
Moderated mediation occurs when the treatment effect of an independent variable A on dependent variable C via a mediator variable B differs depending on levels of a moderator variable D. Specifically, either the effect of A on the B, and/or the effect of B on C depends on the level of D.

Question 3

Explain the following statements

a) A mediator must be related to the predictor variable.
   Mediation model proposes that the independent variable influences the mediator variable, which in turn influences the dependent variable. In this we need to regress mediator on the independent variable to verify that the independent variable is a significant predictor of the mediator. If the mediator is not associated with the independent variable, then it couldnt possibly mediate anything.

b) A moderator can create problems with multicollinearity
   In moderated regression analysis, a new interaction predictor (x1, x2) is calculated. However, the new interaction term will be correlated with the two main effects terms used to calculate it. This is the problem of multicollinearity in moderated regression. Multicollinearity tends to cause coefficients to be estimated with higher standard errors and hence greater uncertainty.

Question 4

Treatment A cures disease B more effectively in males than in females. Furthermore, treatment A cures disease B by lowering overall blood volume. What is the moderating variable in this scenario? What is the mediating variable in this scenario? Justify your answers briefly.

Solution:
The moderating variable is the lower overall blood volume. This is because treatment A leads to an decrease in blood volume which in turn is curing the disease B. So Blood volume is mediating over the variable treatment A which has an effect on Disease B.

Question 5

A researcher wants to know if the self-reported anxiety and happiness of Research Methods II students predicts how much coffee they drink a day. Using the Homework 3 dataset regress coffee consumption on happiness and anxiety scores. Discuss statistical significance and interpret what the constant and each regression coefficient means in relation to coffee consumption.

Solution:
Multiple regression analysis was performed to determine how well a combination of anxiety and happiness scores predicted the coffee drinks a day. The regression analysis revealed that a combination of the variables happiness and anxiety accounted for a significant amount of variance in the variable coffee intake a day, $R^2 = .6, F(2,27) = 20.27, p < .001$, at the .05 level of significance. The regression analysis also revealed that within the model, the variable anxiety accounted for a significant amount of variance in coffee intake, $b = 1.06, t(27) = 4.89, p < 0.001$, however the variable anxiety didn't account for a significant variance in "Coffee Intake", $b = -0.37, t(27) = -1.88, p = 0.07. Alpha = .05$. In the absence of anxiety and happiness, the coffee intake was reported as the constant, $b = 2.33, (27)t = 2.24, p = 0.033$.

Question 6

Reconduct the analysis in problem 5 using grand-mean centered variables. Discuss statistical significance and interpret what the constant and each regression coefficient means in relation to coffee consumption.

Solution:
Multiple regression analysis was performed to determine how well a combination of anxiety and happiness scores predicted the coffee drinks a day. The regression analysis revealed that a combination of the variables happiness and anxiety accounted for a significant amount of variance in the variable coffee intake a day, $R^2 = .6, F(2,27) = 20.27, p < .001$, at the .05 level of significance. The regression analysis also revealed that within the model, the variable anxiety accounted for a significant amount of variance in coffee intake, $b = 1.06, t(27) = 4.89, p < 0.001$, however the variable anxiety didn't account for a significant variance in "Coffee Intake", $b = -0.37, t(27) = -1.88, p = 0.07. Alpha = .05$. In the absence of anxiety and happiness, the coffee intake was reported as the constant, $b = 4.36, t(27) = 16.91, p < 0.001$.

Question 7

Are your interpretations different in problems 5 and 6? Explain why each interpretation is different or the same.

Solution:
The model fit and the overall R^2 doesn't change with centering. The b value and significance value for main effect also doesn't show any changes in the above two model. But there is a intercept shift (constant value changes) and are now interpretable based on comparison to the mean of the data.

Question 8

Was it appropriate or inappropriate to center the variables? Justify your decision.

Solution:
It was appropriate to center the variables happiness and anxiety, because zero value for these two variables have no meaning. Centering the variables so that the predictors have mean 0. This makes the intercept term to be interpreted as the expected value of outcome variable when the predictor values are set to their means.

Attachments

Output graph of the data into SPSS.