Mock exam "Multivariate Statistics", fall semester 2023

Linear regression

Multicollinearity (Kprim) (Kprim) - 2

Which of the following statements about multicollinearity are correct, and which are incorrect?
Decide whether the following statements are true or false by clicking the respective box.

	True	False
In the presence of multicollinearity, there is an increase in standard errors of the estimates.	☐	☐
Multicollinearity occurs when researchers include many highly correlated variables.	☐	☐
Collinearity statistics include the Tolerance and the VIF.	☐	☐
Theory-driven and intelligent variable selection can help to avoid multicollinearity.	☐	☐

Linear regression (open-ended) (Essay) - 1

What is the added-value of adjusted R-squared compared to single R-squared?
Tips: think about the number of predictors and the number of cases.

Logistic regression

Logistic regression (Kprim) (Kprim) - 2

Which of the following statements about logistic regression are correct, and which are incorrect?
Decide whether the following statements are true or false by clicking the respective box.

	True	False
Logistic regression is used to make predictions about a dichotomous dependent variable.	☐	☐
Odds can be defined as the number of times something occurs relative to the number of times it does not occur.	☐	☐
If the odds ratio of a dummy variable is greater than 1, then the group captured in the dummy variable is predicted to be more likely than the reference group to have something occur.	☐	☐
When there is exactly a 0.5 probability of something occurring, the log odds are 1.	☐	☐

Logistic regression (open-ended) (Essay) - 2

Linear regression is NOT appropriate when the dependent variable is a dichotomous (or binary) variable. Explain why this is the case by referring to two properties of the logistic regression.
Tips: think about properties such as normal distribution, variance of the residual errors, and predicted values.

Moderation/Mediation

Moderation (single choice) (Choice) - 1

Which of the following statements about the interpretation of the moderation effect is correct?
Recall that the regression equation for moderation is: Y = a + b1*X + b2*Z + b3*XZ + e = a + b2*Z + (b1 + b3Z)*X + e
Decide which of the following statements is correct by clicking on a single box.

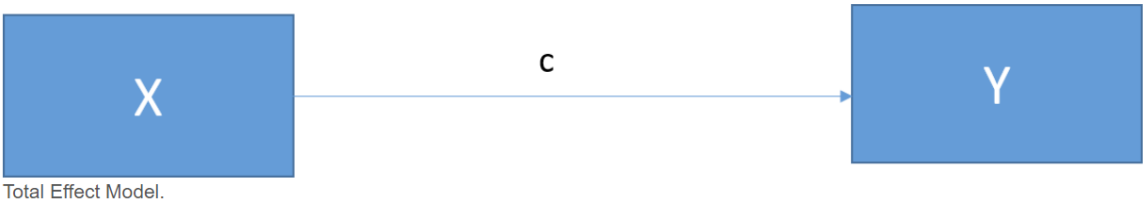
The coefficient b1 can be the main effect of X from a 2X2 ANOVA if X and Z are coded as numeric variables.	☐
The coefficient b1 is the conditional effect of Z on Y when X = 0.	☐
When X = 0, the conditional effect of Z reduces to b3.	☐
The coefficient b1 is the estimated difference in Y between two cases in the data that differ by one unit in X but have a value of 0 for Z.	☐

Moderation/Mediation (open-ended) (Essay) - 2

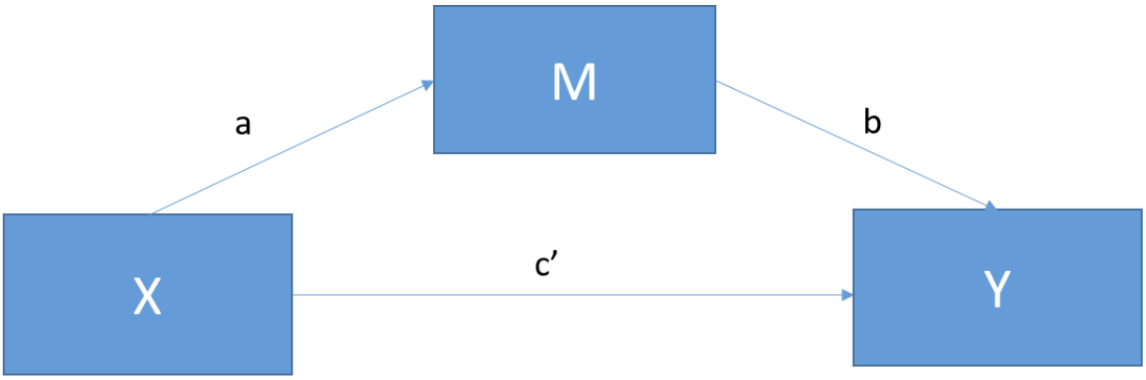
What is conceptionally the difference between moderation analysis and mediation analysis?

Mediation (single choice) (Choice) - 1

Which of the following statements about mediation analysis is NOT correct?



Total Effect Model.



Basic Mediation Model.

Decide which of the following statements is NOT correct by clicking on a single box.

Mediation analysis should only be performed when this total effect of X, path c, is statistically different from zero	☐
In mediation analysis, the size of ab is not determined by c or c'.	☐
The Sobel test has been criticized because it assumes the sampling distribution of ab is normal in form	☐
In mediation analysis, ab could be large even though c is small	☐

ANOVA

Repeated Measurement ANOVA (single choice) (Choice) - 1

Which of the following about the partitioning of the error term in repeated measures (RM) ANOVA is correct?
Decide which of the following statements is correct by clicking on a single box.

RM ANOVA does not further partition the error term from ANOVA.	☐
The partitioning of the within-group variability in RM ANOVA decreases the value of the F-statistic.	☐
The partitioning of the within-group variability in RM ANOVA reduces the size of the error term compared to ANOVA.	☐
The partitioning of the within-group variability in RM ANOVA decreases the power of the test to detect significant differences between means.	☐

ANOVA (open-ended) (Essay) - 3

Mention and describe the three components of the variance in ANOVA.

ANOVA (Kprim) (Kprim) - 2

When are post hoc tests useful in one-way ANOVA?
Tips: recall that one-way ANOVA is when there is one independent variable and one dependent variable.
Decide whether the following statements are true or false by clicking the respective box.

	True	False
Post hoc tests are useful when a factor has more than two levels.	☐	☐
Post hoc tests are useful when differences between individual factor levels from two different factors are to be tested.	☐	☐
Post hoc tests are useful when differences between individual factor levels from the same factor are to be tested.	☐	☐
Post hoc tests are useful when the p-value associated with the F-test is statistically significant.	☐	☐

Factor analysis (EFA/CFA)

EFA (single choice) (Choice) - 1

Which of the following statement about exploratory factor analysis (EFA) is correct?
Decide which of the following statements is correct by clicking on a single box.

The KMO (Kaiser-Meyer-Olkin measure of sampling adequacy) describes the amount of variance of one item that is explained by all factors.	☐
The Kaiser's criterion is a measure of whether the data is suitable for an exploratory factor analysis.	☐
In EFA each variable cannot be fully explained by a linear combination of the factors.	☐
Bartlett's test has a significant value when correlations between items are not large enough to be used in factor analysis.	☐

CFA (open-ended) (Essay) - 1

Name two kinds of parameters in a CFA and define them.

CFA (single choice) (Choice) - 1

In confirmatory factor analysis (CFA), which of the following statement is NOT a mathematical requirement?
Decide which of the following statements is NOT correct by clicking on a single box.

Indicators should not correlate with one another.	☐
There are several interval-scaled items, each of which is normally distributed.	☐
Items that should theoretically load on a factor should not correlate empirically.	☐
Sufficient directly measured items must be available in order to be able to test the assumed model structure made up of items and factors.	☐

SEM

SEM (Kprim) (Kprim) - 2

Which of the following statements about structural equation modelling (SEM) are correct, and which are incorrect?
Decide whether the following statements are true or false by clicking the respective box.

	True	False
SEM is estimated so as to minimize the difference between the observed and estimated covariance matrices	☐	☐
SEM is estimated so as to minimize the sum of squares between constructs	☐	☐
It is recommended that each construct be measures by at least three observed variables	☐	☐
Confirmatory factor analysis is a type of SEM	☐	☐

MLM

MLM (Kprim) (Kprim) - 2

Imagine you have the following two-level null model specified for a dependent variable Yij (i: students and j: schools):

- Level 1 (students): Yij = b0j + rij
- Level 2 (schools): b0j = g00 + u0j

Which of the following statements are correct, and which are incorrect?
Decide whether the following statements are true or false by clicking the respective box.

	True	False
b0j represents the group-mean of Y	☐	☐
rij represents level-2 residuals	☐	☐
g00 represents the grand mean of the variable Y	☐	☐
u0j represents the group-specific deviation from the grand mean of Y	☐	☐

Bonus

Bonus (open-ended) (Essay) - 2

Imagine that survey items investigating which factors are related to COVID-19 vaccine hesitancy for adults have been collected and that you are in charge of making a report. Which method(s) of analysis would you use and why?