

MITx: 14.310x Data Analysis for Social Scientists

Heli



 Module 1: The Basics of R and Introduction to the Course

- ▶ Entrance Survey
- Module 2: Fundamentals of Probability, Random Variables, Distributions, and Joint Distributions
- Module 3: Gathering and Collecting Data, Ethics, and Kernel Density Estimates
- Module 4: Joint,
 Marginal, and
 Conditional
 Distributions &
 Functions of Random
 Variable

Module 9: Single and Multivariate Linear Models > The Multivariate Linear Model > An Example of the Multivariate Linear Model - Quiz

An Example of the Multivariate Linear Model - Quiz

☐ Bookmark this page

Question 1

1/1 point (graded)

You are interested in finding out what affects workers' happiness. To this goal, you administer a survey, your outcome is self-reported happiness. You think that time spent at work and free time are important predictors of workers' happiness, so you also ask workers to report the number of hours they spend at work and the number of hours they have outside of work. To see which of those variables is a better predictor of happiness, you regress self-reported happiness on both these variables and a number of other variables you collected data on (age, gender, income).

Which of the following (if any) is definitely going to be a problem with this regression?

- 🏿 a. There is a collinearity problem. 🗸
- b. The errors are correlated across observations.
- c. You don't have enough positive variation in the regressors.
- d. There is no problem with this regression.

- Module 5: Moments of a Random Variable,
 Applications to Auctions,
 Intro to Regression
- Module 6: Special
 Distributions, the
 Sample Mean, the
 Central Limit Theorem,
 and Estimation
- Module 7: Assessing and Deriving Estimators -Confidence Intervals, and Hypothesis Testing
- Module 8: Causality,
 Analyzing Randomized
 Experiments, &
 Nonparametric
 Regression
- Module 9: Single and <u>Multivariate Linear</u> <u>Models</u>

The Linear Model
due Nov 28, 2016 05:00 IST

Explanation

The number of hours spent working is by definition 24 minus the number of hours spent outside of work. So your regressors are linearly dependent, and you have a collinearity problem. Intuitively, this is going to be a problem because we don't have enough variation to identify the effects of these 2 variables because they are perfectly collinear (basically just flip sides of the same thing).

Submit

You have used 1 of 2 attempts

Correct (1/1 point)

Question 2

1/1 point (graded)

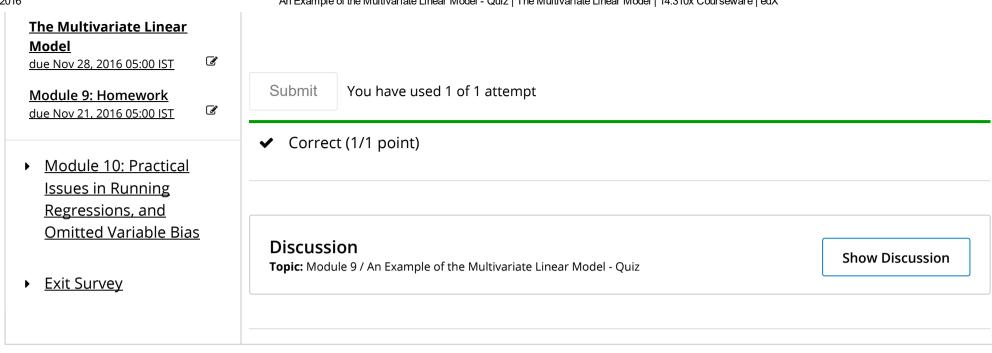
True or False: Consider the model $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \epsilon$. If 2 of your regressors are closely related in the following way $X_3 \approx X_2 + 3X_1$, but not perfectly, your model is still estimable.

a. True

b. False

Explanation

As Prof. Ellison explained in lecture, your model is still estimable, however the resulting estimator will have a very high variance. Depending on the context, you may decide that dropping one of these variables makes sense in order for you to obtain estimates with reasonable variance. However, in that case, your estimates might be slightly biased.



© All Rights Reserved



© 2016 edX Inc. All rights reserved except where noted. EdX, Open edX and the edX and Open EdX logos are registered trademarks or trademarks of edX Inc.















