Finger Exercises due Nov 2, 2020 18:30 EST Past Due

Question 1

0.0/1.0 point (graded)

Which of the following are assumptions of the Multivariate Linear Model? Select all that apply.

- [ ] $E[\epsilon] = 0$ ✔
- [ ] The number of observations is greater than the number of regressors. ✔
- [ ] $Cov(Y, \epsilon) = 0$
- [ ] The errors are uncorrelated across observations. ✔
- [ ] The regressors are linearly independent. ✔

Explanation

These are all assumptions of the Multivariate Linear Model except $Cov(Y, \epsilon) = 0$, which is not an assumption because we are not making claims about causality.

Show answer

Question 1

0.0/1.0 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?

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

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).

Question 2

0.0/1.0 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.

- ◯ True ✔
- ◯ False

Explanation

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.

Show answer

Submit You have used 0 of 1 attempt

Answers are displayed within the problem

The Multivariate Linear Model Continued... - Quiz

Bookmark this page

Finger Exercises due Nov 2, 2020 18:30 EST **Past Due**

Question 1

0.0/1.0 point (graded)

True or False: Suppose you are running a regression, and are worried that two of your variables might be perfectly collinear. You start by looking at the correlation coefficients, and find that they are highly correlated $\rho = 0.9$. This means that they are perfectly collinear.

- True
- False ✔

Explanation

A high correlation coefficient implies your variables are strongly related to each other, however it doesn't imply that your variables are perfectly collinear, because they maybe related by some non-linear relationship.

Show answer

Submit You have used 0 of 1 attempt

Question 2

0.0/1.0 point (graded)

Suppose you are estimating a model $Y = \alpha + \beta X + \epsilon$, where $\epsilon = \begin{bmatrix} \epsilon_1 \\ \vdots \\ \epsilon_n \end{bmatrix}$ Considering each of the following statements independently. Which of the following would imply that $E\left[\epsilon\epsilon^T\right] \neq \sigma^2 I$? (Select all that apply)

- [ ] $\mathrm{Cov}\left[\epsilon_i \epsilon_j\right] \neq 0$ for some $i, j \in (1, \ldots, n)\,,\ i \neq j$ ✔
- [ ] $\mathrm{Cov}\left[\epsilon_i \epsilon_j\right] \neq 0$ for all $i, j \in (1, \ldots, n)\,,\ i \neq j$ ✔
- [ ] $\mathrm{Cov}\left(\epsilon\right) \neq \sigma^2 I$ ✔
- [ ] Your errors are correlated across observations ✔

Explanation

They're all true! First, recall that $E\left[\epsilon\epsilon^T\right]$ is sometimes denoted as $\mathrm{Cov}\left(\epsilon\right)$ (the variance-covariance matrix of ϵ). So C is equivalent to $E\left[\epsilon\epsilon^T\right] \neq \sigma^2 I$. The off-diagonal elements of the n x n matrix $E\left[\epsilon\epsilon^T\right]$ are given by $\mathrm{Var}\left(\epsilon_i\right)$ for all $i, j \in\ (1, \ldots, n)\,, i \neq j$. Whereas the elements on the diagonal are $\mathrm{Var}\left(\epsilon_i\right)$ for all $i \in (1, \ldots, n)$. So, in order for $E\left[\epsilon\epsilon^T\right] = \sigma^2 I$, all the off-diagonal elements must be equal to 0, i.e. $\mathrm{Cov}\left(\epsilon_i \epsilon_j\right) = 0$ for all $i, j \in (1, \ldots, n)\,,\ i \neq j$, or in words: the errors must be uncorrelated across observations.

Deriving Estimators in MV Linear Model - Quiz

Bookmark this page

Finger Exercises due Nov 2, 2020 18:30 EST **Past Due**

Question 1

0.0/1.0 point (graded)

True or False: If the errors are not normally distributed, $E\left[\hat{\beta}\right] \neq \beta$.

- True
- False ✔

Explanation

The distribution of the errors only affects the distribution of $\hat{\beta}$. However, $\hat{\beta}$ is an unbiased estimator irrespective of the distribution of the errors. This assumption is useful for inference purposes (recall the discussion on hypothesis testing, you need to make an assumption about the underlying distribution so you have something to compare your estimator to!).

Show answer

Submit You have used 0 of 1 attempt

What happens to your estimator for the variance as you add more regressors? (Select all that apply)

- [ ] The variance decreases if variable has some explanatory power. ✔
- [ ] The variance increases if variable has no explanatory power. ✔
- [ ] The variance decreases if variable has no explanatory power.
- [ ] The variance increases if variable has some explanatory power.
- [ ] This does not affect variance.

Explanation

Recall the formula for the variance estimator:

$$\hat{\sigma}^2 = \frac{\hat{\epsilon}^T\hat{\epsilon}}{(n-k)}$$

Where $\hat{\epsilon}^T\hat{\epsilon}$ denotes the residuals, and $\hat{\sigma}^2$ denotes the residual variance (i.e the variance of that's not explained by your regressors.) So if your additional regressor has no explanatory power, adding a regressor will just increase your degrees of freedom k, and leave the residual variance fixed, thereby increasing your variance estimator. However, if your regressor

Suppose you are interested in testing hypotheses of the following form:

$$H_0 : R\beta = c$$

$$H_1 : R\beta \neq c$$

where β is a column vector.

Which of the following statements is true about the matrix R?

- ◯ The number of rows in R corresponds to the number of restrictions you want to test, and the number of columns corresponds to the number of parameters in your model. ✔
- ◯ R is an invertible square matrix.
- ◯ R is symmetric
- ◯ The number of rows in R corresponds to the number parameters in your model, and the number of columns corresponds to the number of restrictions you want to test.

Explanation

R is a matrix of restrictions with dimensions r x $(k+1)$, where r is the number of restrictions you want to test, and $k+1$ is the number of parameters in your model.

Finger Exercises due Nov 2, 2020 18:30 EST Past Due

Question 1

0.0/1.0 point (graded)

Suppose you are interested in testing hypotheses of the following form:

$$H_0 : R\beta = c$$

$$H_1 : R\beta \neq c$$

where elements of β are $\beta_0, \beta_1, \ldots, \beta_5$

Suppose R and c are as follows:

$$R = [0\ 0\ 1\ 1\ 1\ 0]\,, c = 1$$

What hypotheses would this be testing?

- ◯ $\beta_2 + \beta_3 + \beta_4 = 1$ ✔
- ◯ The set of coefficients $\beta_1 = \beta_2 = \beta_3 = 1,\ \beta_4 = 0$
- ◯ The set of coefficients $\beta_2 = \beta_3 = \beta_4 = 1,\ \beta_5 = 0$
- ◯ A subset of the coefficients are all equal.

Explanation

Write out the matrix multiplication:

$$R\beta = [0\ 0\ 1\ 1\ 1\ 0]\begin{bmatrix}\beta_0\\ \beta_1\\ \beta_2\\ \beta_3\\ \beta_4\\ \beta_5\end{bmatrix} = \beta_2 + \beta_3 + \beta_4 = c = 1$$

Show answer

Submit

You have used 0 of 2 attempts

Answers are displayed within the problem

Question 2

0.0/8.0 points (graded)

Using the same setup discussed in class, and in the previous questions, suppose we have a model with 5 parameters, and want to test the hypotheses that $\beta_1 = \beta_2, \beta_3 + \beta_4 = 1$, and $\beta_5 = 10$. Fill in the missing elements (A, B, C, D, E, F, G, H) of the matrix of restrictions R and the vector c:

$$R\beta = \begin{bmatrix} 0 & 1 & A & 0 & 0 & 0 \\ 0 & 0 & B & C & 1 & 0 \\ 0 & 0 & 0 & 0 & D & E \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_5 \end{bmatrix} = c = \begin{bmatrix} F \\ G \\ H \end{bmatrix}$$

A:

Answer: -1

B:

Answer: 0

C:

Answer: 1

D:

Answer: 0

E:

Answer: 1

F:

Answer: 0

G:

Answer: 1

H:

H:

Answer: 10

Explanation

$$R\beta = \begin{bmatrix} 0 & 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_5 \end{bmatrix} = c = \begin{bmatrix} 0 \\ 1 \\ 10 \end{bmatrix}$$

This follows immediately from the matrix multiplication

$$R\beta = \begin{bmatrix} 0 & 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_5 \end{bmatrix} = \begin{bmatrix} \beta_1 - \beta_2 \\ \beta_3 + \beta_4 \\ \beta_5 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 10 \end{bmatrix}$$

Show answer

Question 1

0.0/1.0 point (graded)

Suppose you have a model with 5 parameters. You are interested in testing the restriction $\beta_1 = \beta_2 = 0$. Which of the following correspond to the restricted and unrestricted models you would compare to test these hypotheses?

- **Restricted:** $Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon$

  Unrestricted: $Y = \beta_0 + \beta_3 x_3 + \beta_4 x_4 + \epsilon$

- **Restricted:** $Y = \beta_0 + \beta_3 x_3 + \beta_4 x_4 + \epsilon$

  Unrestricted: $Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \epsilon$

  ✔

- **Restricted** $Y = \beta_0 + \beta_3 x_3 + \beta_4 x_4 + \epsilon$

  Unrestricted: $Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon$

- **Restricted:** $Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon$

  Unrestricted: $Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \epsilon$

Explanation

The unrestricted model is the model imposing no restrictions on any of the parameters. So C and D are clearly wrong. In order to test that the coefficients on X_1, X_2 are both 0, you want to see whether they affect the explanatory variable, intuitively if the coefficients are 0, you will find that excluding them makes no difference on the goodness of fit. To that goal, your restricted model is the model where $X_1, X_2$are omitted. So B is the correct answer.

Show answer

Submit You have used 0 of 2 attempts

Answers are displayed within the problem

Question 2

0.0/1.0 point (graded)

Continuing with the same example, what is the alternative hypothesis H_1 if you want to test that hypothesis using an F test?

- $\beta_1 > 0, \beta_2 > 0$
- $\beta_1 \neq \beta_2 \neq 0$
- At least one of the coefficients is different from 0. ✔

◯ Both of the coefficients are different from 0.

Explanation

The best way to think of the alternative hypothesis is to think what is necessarily true if the null is false? In this case, the null is not true implies either $\beta_1 \neq 0$ or $\beta_2 \neq 0$ or both. However, if you just reject the null, you know that one of these must be true, but it doesn't necessarily imply that both are not true.

Show answer

Submit You have used 0 of 2 attempts

Answers are displayed within the problem

Question 3

0.0/1.0 point (graded)

Continuing with the same example. Assume you have 500 observations, and suppose you run the regressions corresponding to the restricted and unrestricted models in question 1. Suppose that the you obtain the following values: SSR_R= 198.3 and SSR_U=183.2. If $\alpha = 0.05$, find the critical value for your F statistic.

Note: You can use tables or an online calculator. Please round all answers to the second decimal place, i.e. if your answer is 0.005, please round to 0.01 and if it is 0.013, please round to 0.01

Answer: 3.01

Explanation
The critical value for your F statistic will depend on $\alpha = 0.05$, the degrees of freedom in the numerator, and the degrees of freedom in the denominator. The degrees of freedom in the numerator are given by the number of restrictions $(r = 2)$. And then the degrees of freedom in the numerator $(n - (k + 1) = 500 - 5 = 495)$. So you can look up $F_{2,495}$ for $\alpha = 0.05$ and you get 3.01

Based on the information above, compute the test statistic.

Round all answers to the first decimal place, i.e. if your answer is 1.193, round to 1.2, and if it is 1.113, round to 1.1

Answer: 20.4

Explanation
As Prof. Ellison discussed in lecture, this is given by: $T = ((SSR_R - SSR_U)/r)/(SSR_U/(n - (k + 1))$. Substituting in the relevent values: $T = ((198.3 - 183.2)/2)/(183.2/495) = 20.4$. Note: That we are conducting the F test, however, to be consistent with the notation used in cl,ass we use T as the notation for the F-test statistic.

Show answer

Question 4

0.0/1.0 point (graded)

Based on the values from Question 3, do you accept or reject the null at the 95% confidence level?

- Accept
- Reject ✔
- Uncertain

Explanation

Under the null the test statistic follows an F-distribution with parameters $(r, n-(k+1))$. So we reject the null for large values of the test statistic. In this case, $T = 20.4$ and $F_{crit} = 3.01$, so $T > F$, and thus we reject the null.

Show answer

Submit You have used 0 of 1 attempt

Answers are displayed within the problem