1

Eco 213 (Data Analysis and Basic Econometrics): Monsoon 2022

**Instructor: Gitanjali Sen Total Marks: 20**

**Take-Home (Graded) Assignment: 9th Nov 2022**  **Due Date: 4th December 2022**

VIKAS\_KUMAR\_GUPTA \_2010110715

**Exercise:**

1. **C4 of Wooldridge** (Computer based exercise from Chapter 3): There are 5 parts in the question. Just to ensure that you get the same question that I want, I am copying the 1st part of the question below. **But you need to do all the parts from 1 to 5 in this question.**

Use the data in ATTEND.RAW for this exercise.

(i) Obtain the minimum, maximum, and average values for the variables atndrte, priGPA, and ACT.

Do all the parts from i) to v).

ANS(1.i):

Observation:680

|  |  |  |  |
| --- | --- | --- | --- |
| Varible | Min | Max | Mean |
| atndrte | 6.25 | 100 | 81.70956 |
| priGPA | .857 | 3.93 | 2.586775 |
| ACT | 13 | 32 | 22.51029 |

ANS(1.ii):

atndrte = β0 + β1priGPA + β2ACT + u,

estimated equation: atndrte = 75.7004 + 17.26059\*priGPA -1.716553\*ACT

N = 680

R-Squared=0.2906

the inercept of 75.7004 is the predicted percent of classes attended for a student with “0”

cumulative GPA prior to the current term and an ACT score of “0”.

No, It doesn’t have a useful meaning because the intercept is useful but its iterpretation is not.

ANS(1.iii):

# increase of 1 GPA is predicted to increase the percentage of classes attended by 17.26059 percent.

This makes a resonable amount of sense 🡪 increasing GPA increases attendance.

# increase of 1 ACT exam, is predicted to decrease attendance. This is unexpected. Accordign to my interpretation , increase in ACT score may indicate an increase in ability and with increased ability it become less necessary to attend classes. I don’t like this story- I would call it a reach at best- but it is consists with the data.

ANS(1.iv):

PriGPA = 3.65

ACT =20

Predicted attendent rate(atndrte) = 75.7004 + 17.26059\*(3.65) + (-1.716553)\*20

= 75.7004 + 63.0011535 -34.33106

=104.3704935%

A student with a GPA of 3.65 and an ACT of 20 would seem to be a very good student. But no student can attends more than 100% of classes!

There is one student with these exact values and there are 2 student meet this criteria(160,569)observation.

Observation(569) has an exact ACT of 20.

ANS(1.v):

Student A: priGPA = 3.1

ACT = 21

A(attendence rate) = 75.7004 + 17.26059\*(3.1) + (-1.716553)\*21

= 75.7004 + 53.507829 - 36.047613

= 93.160616

Student B: priGPA = 2.1

ACT = 26

B(attendence rate) = 75.7004 + 17.26059\*(2.1) + (-1.716553)\*26

= 75.7004 + 36.247239 - 44.630378

= 93.160616

Predicted attendence rate difference = (93.160616-93.160616)%=25.843355 %

**12 marks**

2. **C8 Wooldridge** (same chapter as above)

Use the data in DISCRIM.RAW to answer this question. These are zip code–level data on prices for various items at fast-food restaurants, along with characteristics of the zipcode population, in New

2

Jersey and Pennsylvania. The idea is to see whether fast-food restaurants charge higher prices in areas with a larger concentration of blacks.

**Do questions (i) to (iii) only.** **8 Marks**

**ANS(2.i):**

|  |  |  |
| --- | --- | --- |
| **variable** | **mean** | **Standard deviation** |
| **prpblock** | **.1134864** | **.1824165** |
| **income** | **47053.78** | **13179.29** |

It is apparent that prpblck represents a proportion of the black population, while income is represented in dollar terms.

ANS(2.ii):

*psoda* = *β*0 + *β*1\**prpblck* + *β*2\**income* + *u.*

psoda = .1855321 + .1258267\*prpblck + .0788228\*income

R-squared = 0.0663 N=401

The coefficient on prpblck is 0.1149882. The literal interpretation would be: when prpblck increases by 1, the price of a medium soda increases by 11.49882 cents. The only problem is, the notion of increasing prpblck by 1 is not very meaningful. prpblck is the proportion of individuals in a zip code who are black cannot increase by 1 unless the proportion of individuals in a zip code starts out as 0. That is, the only zip code that can increase by 1 is a zip code that starts out with no individuals who are black, and then becomes a zip code that is made up only of individuals who are black. This is not a very useful marginal effect.

In order to interpret the marginal effect more usefully, look at smaller (more realistically-sized) changes. For instance, an increase of 0.01 (an increase of 1 in the percentage of individuals who are black in a zip code) is predicted to increase

the price of a medium soda by 0.1149882 *×* 0.01 = 0.00114988, or approximately not at all. An increase of 0.10 (an increase of 10 in the percentage of individuals who are black in a zip code) is predicted to increase the price of a medium soda by 0.1149882 *×*

0.10 = 0.0114988, or approximately a penny. I do not think it is economically large because

there are many medium sodas purchased, such an effect might be better expressed in terms of total expenditure, which might be large. A penny, multiplied by many tens of thousands of sodas in a particular zip code, starts to run into real money.

**ANS(2.iii):**

Result of a simple regression of psoda on prpblck:

*psoda* = 1.037399+ .0649269\**prpblck*

ˆ

*N* = 401 *R − squared* = 0.0181

The discrimination effect is estimated to be significantly smaller when income is excluded from the regression.