Problem Set 2

Q 1 A/B Test

In this question, you will be working on a dataset, **abtest.csv**, from an A/B test on an e-commerce website. Consumers are randomly assigned to old page or new page. The variable of interest, **converted**, is the binary outcome variable.

a) How many consumers see the new page? What is the conversion rate for new page? How many consumers see with the old page? What is the conversion rate for old page?

b) What is the proportion of consumers converting regardless of the page they receive?

c) Conduct a A/B test using two-sample test for proportion mentioned in class. Please state your null hypothesis, report associated z-statistics, and comment on your result.

Q 2 Demand Estimation: Homogeneous Product

Use **demandIV.dta** for this question. Consider simple demand and supply equations for a homogeneous product:

$$quantity = \beta_0 + \beta_1 price + \beta_2 pcompete + \beta_3 income + \epsilon_1$$

$$price = \alpha_0 + \alpha_1 quantity + \alpha_2 praw + \epsilon_2$$

where *quantity* is the quantity of a product produced and sold, *price* is the price of the product, *pcompete* is the price of a competing product, *income* is the average income level of consumers, and *praw* is the price of raw materials used to produce the product.

a) Please estimate the equations using OLS.

b) Please estimate the equations using 2SLS. You use demand shifters as IV for supply equation, and supply shifter as IV for demand equation.

c) Compare the results and give your comments.

Q 3 Demand Estimation: Berry-Logit

This question uses **otc.dta** on "Over the counter" Headache medicine (i.e. aspirine, tylenol and such). We have data on the number of customers, product sales, retail prices, wholesale prices (the retailer's costs) in each store and in each week.

- Brand: an ID for different medicine (Brand dummies are also provided in the data-set)

- Sales: total sales for that brand at the store each week
- Count: total number of people that go into the store each week
- Price: Price of the medicine
- Promotion: promotion on the product in the store each week
- Cost: wholesale prices (the retailer's costs)
- Average Price: the average price of other products in other stores in the same week
- Week and Store are the time and market indicator

Consider the utility function for product j in store-week t for consumer i:

$$u_{ijt} = \boldsymbol{X}_{jt}\boldsymbol{\beta} - \alpha p_{jt} + \xi_{jt} + \epsilon_{ijt} = \delta_{jt} + \epsilon_{ijt} \tag{1}$$

where ϵ_{ijt} is an i.i.d. logit draw, $\boldsymbol{X}_{jt}$ are observed product characteristics, ξ_{jt} are unobserved product characteristics, and δ_{jt} is the mean utility from product j in week t.

Berry (1994) shows how to analyze this model by solving for δ as a function of observed market shares, obtaining

$$\delta_{jt} = \ln(s_{jt}) - \ln(s_{0t}) = \boldsymbol{X}_{jt}\boldsymbol{\beta} - \alpha p_{jt} + \xi_{jt}$$

This formulation allows us to consider the linear regression model to estimate demand.

a) You have to first calculate the market shares for each store and in each week. Total number of people that go into the store each week is defined as the whole population. Outside option is not purchasing any of the medicine in the store. Please estimate this model using different specifications:

   (i) Run an OLS regression with price and promotion as product characteristics.

   (ii) Using the analytic formula for elasticity of the logit model and price coefficient in the previous OLS regression, compute the own-price elasticities for all observations and plot a histogram. What are the mean own-price elasticities for all 11 brands?. Do these results for elasticities make sense? Why, or why not?

   (iii) Use **cost_** in 2SLS as an instrument to price, compare with the results with previous OLS regression. What do you find?

   (iv) (Bonus Question) Suppose that you are interested in the price sensitivities from consumers, and you are not satisfy with the results from previous results, please propose a solution and implement it to see if you get more reasonable price elasticities.