**Economics of Discontinuous Change 1818: 3/4s Exam – April 2020**

FORMAT & INSTRUCTIONS:

Open book. You may refer to your notes and readings. **EXPLAIN** clearly and concisely in your own words.

If there is confusion about a question, please query either freeman@nber.org or yangyou@g.harvard.edu

Please be sure to **insert your name in the HEADER above.**

Return exam to by Saturday 4/11 @ midnight EDT:
Send to: amadeoholl@gmail.com
E-mail subject line: LAST NAME, First Name\_ec1818-3/4ths exam
File Format: MS-word or PDF

File Name: LAST NAME-First Name\_ec1818-mid-exam\_DATE

**QUESTION 1:**You are at the AI Venture Capitalist Ball, and chatting with venture capitalist robots (VCR) who may put money into your new start-up. You would like to strike the best deal, but you know that, once you move on to chatting with another robot, the VCR you talked with will invest in some other project. Thus, you have to decide on the spot whether to accept the offer from the robot you are chatting with.

a) How should you maximize your chances of accepting the best offer (assuming a best offer exists)?

b) What information would you need for this strategy?

c) Your start-up partner says, “Your decision process sounds like numerology or alchemy. Why don't we just flip a coin.” Explain the logic behind your strategy.

d) If you follow this strategy, what is the chance you would actually get the best offer?

**QUESTION 2:**The leader of your organization tweets to customers regularly and rates the tweets by how many get retweeted. Here is data on today's 12 tweets

 tweet # Retweeted

 1 62

 2 31

 3 24

 4 18

 5-6tie 15

 7 14

 8-9tie 12

 10-12 tie 11

 An economics student says, the retweets fit a Pareto distribution and organizes the data like this: number of tweets with over 30 retweets: 2

 number of tweets with over 23 retweets: 3

 number of tweets with over 10 retweets: 12

 This proves the tweets are Pareto distributed and not that silly Zipf law

a) What is the Pareto distribution and what data do economists find usually fits it?

b) What is Zipf's law and what data does it fit?

A physics student say, Pareto-Tomato Zipf-Pf. I know more than you do, and you know it. This is clearly a power law.

c) Does the data fit a power law? How would you estimate the power law’s exponent (alpha)?

d) Professor Zipf rises from the grave and says, “Wait a second, dingbats. The data looks Zippy to me. Just look ...” Fill in what Professor Zipf should say to claim the data. Follow Zipf's law.

e) What does this disagreement tell you about analyzing distributions with fat tails?

**QUESTION 3:**

One of your classmates says, my class paper will examine network links between people who use Pareto curves, Zipf curves, and Power laws. I will download journal articles who say Pareto, and journal articles who say Zipf, and journal articles who say power law and see if there are links between the **co-authors** of Pareto-papers, Zipf-papers, and power law papers.

a) Using network analysis/graph theory of nodes and edges, how would you organize this data?

b) If the co-author links fit the small network model, what would the data show?

Another student says, my class paper links **citations** between papers that use Pareto curves, Zipf Curves, and Power laws.

c) Using network analysis/graph theory of nodes and edges, how would you organize this data?

d) What, if any, differences would you expect from the citation network than from the co-author network?

**QUESTION 4:**

You want to go to the weekly Tuba Skinny concert in Cambridge Common but only if you have 6 feet of space around you for social distancing. Everyone else has the same view. You go if it is not too crowded. You have information about whether the number at the concert in the past 3 weeks was too crowded or met your requirements.

Analyze the choice using the minority game framework.

a) How do people make decisions in the minority game model?

b) If a person flips a coin to go to a concert, will this produce a more or less efficient outcome than if they make decisions along the lines posited in the minority game?

c) How might expanding the number of past weeks over which people make their decision to say 10 weeks affect the outcomes?

**QUESTION 5:**

Algorithm Search for Best Solutions

a) A computer algorithm is searching for value X\* that minimizes the cost of production C(X). The landscape of costs is a rugged one. Suppose the program discovers that C(X1 ) < C(X2 ) but still moves from X1 to X2 with positive probability. Does the program have a glitch?

b) The N-K Model builds a "correlated fitness landscape". What is N? What is K? What is a correlated fitness landscape? How does it differ from a rugged landscape?

c) Suppose costs depend on 5 inputs, each of which can take on 4 states. What value of K would lead to a single peak landscape? What value of K would lead to a completely random landscape?

**QUESTION 6:**

Identify and with a sentence or two and explain their significance:

a.) Fuzzy sets

b.) Minimum stabilizing frequency

c.) Black swan event

d.) Cellular Automata

e.) Zero-determinant Dyson-Press strategy

**QUESTION 7:**

Consider the repeated prisoner’s dilemma game with two rounds and discount factor w (you may interpret this game as one of unknown duration with an expected length of two rounds):

 The PD Game

|  |  |  |
| --- | --- | --- |
|  |  | *Player Two* |
|  |  | *Cooperate* | *Defect* |
| *Player* | *Cooperate* | *3,3* | *0,5* |
| *One* | *Defect* | *5,0* | *1,1* |
|  |  |  |  |

a)    Determine the values of w for TFT to be an equilibrium strategy.

b)    Consider a population of N players that play this game against each other. What are the equilibrium strategy compositions of the population depending on the value of w? Consider only the strategies C (always cooperate), D (always defect) and TFT.

c)    Show that TFT is not an ESS of this game.

**QUESTION 8:** Big Data

a) What is the **over-fitting problem** in machine learning/fitting a model with big data sets?

Illustrate the problem in the case of tree models.

b) How do you resolve the over-fitting problem?

c) The Get Rich Quick investment firm says the best way to predict profitability of new investments is to do an ensemble analysis. It estimates different models – from linear regression to trees to neural nets to logistic equations that link economic data and number of tweets to whether a past investment was profitable (=1) or not (=0). It applies the estimated models to predict whether possible new investments are likely to be profitable or not and invests in the projects getting the most votes. Why might this voting procedure work?

**QUESTION 9: Big thoughts about class material**

a) Neural net models are global models while tree models are local models. Explain.

b) Econ 1818 is about rare discontinuous events that affect economies/lives much more than frequent small events, but we only have a few observations on rare events? How can big data illuminate rare events?