Stochastic Decision Theory (2MMS50)

Assignment I

Please provide brief but careful motivations of your answers.

A superhuman has contacted you to help them invest for their pension. This superhuman invests in two types of economic instruments. One is a cash account that delivers the risk-free interest rate at time t denoted by $R_{f,t}$, and the other is a stock with return $R_{r,t}$ at time t. Note, that the indices mean that over the period from t to $t+1$, one has a return of $R_{\cdot,t}$. One of your goals as investor is to find the fraction of wealth you invest in stocks, denoted by ω_t.

The underlying dynamics of the of the risky asset return is given by

$$R_{r,t} = \begin{cases} (1+u)R_{f,t} & \text{w.p. } \frac{1}{2}, \\ (1+d)R_{f,t} & \text{w.p. } \frac{1}{2}. \end{cases}$$

With $u = 0.8$, $d = -0.3$ and $R_{f,t} = 1.02$.

Furthermore, the individual gets happy when it is allowed to consume some of it's money. So you also have to find C_t which is the amount of money you allow the individual to consume. In order to measure the happiness which the individual receives when consuming C_t, you assume that this behaves according to some utility function $u(\cdot)$, defined as:

$$u(C_t) = \max\{\log C_t, 0\} \tag{1}$$

Furthermore, the individual has a superpower and thus never dies. So you formulate your goal to be, finding ω_t and C_t which maximizes

$$\mathbb{E}_0 \left[\sum_{\tau=0}^{\infty} \delta^{\tau} u(C_t) \right]. \tag{2}$$

Here δ is the subjective discount factor, our individual has chosen for a $\delta = 1$ as it lives forever. As you do not yet know how deal with continuous decision variable you decide to set $\omega_t \in [0,1]$ and discretize this space in to 5 evenly spaced points. Furthermore, you cannot consume more than the wealth you have and you only take integer C_t's. Additionally, you know that this superhuman is not very good in maths. So his wealth W_t you round down to the nearest point in $\{0, 1, \ldots, 10\}$ at the very end of each year. The difference in wealth and the discretized point you pocket as you like to earn some extra money.

The app you use to invest in stocks: BigStocks, knows your famous superhuman client. So you, your client and BigStocks agree that your client promotes their app and in return the superhuman get's a salary s_t at the start of each year. As this salary depends on the popularity of BigStocks it is equal to 1 or 2 both with probability 1/2.

To clarify how the system works; the wealth at time t denoted by W_t is the wealth of the individual at the very beginning of the year. Right after the year starts the individual consumes C_t and receives the s_t unit of money from BigStocks. For the remaining 364 days of the year, you invest the remaining money in the cash-account and the stocks in a $1 - \omega_t : \omega_t$ ratio.

a) Formulate the above problem as a Markov Decision Process. Make sure to include the state space $\mathcal{I}$, action space $\mathcal{A}$, direct rewards $r^a(\cdot)$ and transition probabilities $p^a(\cdot,\cdot)$. In order to find $p^a(\cdot,\cdot)$ you will need a recursive relationship that describes the dynamics of the individual's wealth. Also show that your probability makes sense, i.e. show that

$$\sum_{j \in \mathcal{I}} p^a(i,j) = 1.$$

b) Your current policy is to play risky so you invest $\omega_t = 0.75$ in stocks and as you are a tiny bit scared of your client you let them consume their entire wealth, i.e. $C_t = W_t$.

Call this policy f^{rg}, where the r stand for risky and the g stands for greedy. Compute the long-term average utility (i.e. how happy is your superhuman) for your current policy and also provide the relative rewards. Also describe how you happy you think the superhuman is with this policy.

c) Perform one step of policy iteration on the your current policy and state the improved policy.

d) Use successive approximation to obtain the optimal policy $f^\star$ and give the corresponding maximum consumption. Are there any interesting features in this optimal policy? Can you explain them?

e) As you have now enlighted yourself with the optimal policy f^* you decide to use it. However, despite your best efforts you are afraid that the superhuman will get angry if you do not allow them to consume anything. Hence you want to ensure that the probability of them not consuming anything is less than 0.2. Hence compute

$$\mathbb{P}(C_t^{f^*} = 0 | W_t \text{ in stationarity}).$$

After this formalise the requirement stated above and detail how an optimal policy can be found under this constraint. *You do not have to compute the solution.*

f) Bonus question. Compute the solution to Question e).