Part A – Mini-Essay oriented questions (solve 3, 10 pts each)

1) Explain the mathematical as well as conceptual meaning of the R^2 statistic, and how it should be used for simple and multiple regression analysis as a metric to judge the merit of the regression.

R square statistic is used as a percentage indicator of the difference in the dependent variable that those that are independent elaborate collectively (Saunders, Russell, & Crabb, 2012). The R squared measures the relationship strength that exists between the dependent variable and one's model on a scale of 0- 100% measure. R-squared finds the scatter belonging to the points of data that are fitted on the regression line. It's also termed the coefficient of multiple determination for multiple regression or the coefficient of determination. Higher R-squared values in the same data indicate smaller variance margins between the fitted values and observed data. Generally, R-squared is the dependent variable variation percentage that the linear model elaborates.

R2= Variance elaborated by the model

Accumulative Variance

Multiple regression is a method used to analyze the relationship between one dependent variable and several independent ones. It is undertaken in the following steps model adequacy, building, validation and assumptions. In regression analysis, R2 points out the level of variation of a dependent variable is elaborated by those that are independent

2) Briefly explain the methods of decision making under uncertainty (Maximax, Maximin, Minimax regret) and under which circumstances they may lead to bad decisions.

Decision-making under uncertain circumstances is known as a decision issue. The decision-maker always has an insufficient knowledge to locate the probabilities of occurrence but is aware of different potential nature states. A decision below uncertainty is in which there are numerous unknowns, and there may be no way of information what would possibly happen in the future to affect the result of choice. Various parameters have been cautioned to choose the correct movement route inside the uncertainty's atmosphere. The strategies of choice-making under fact are each of those parameters makes an inference approximately the choice mindset.

Maximax criterion: If the decision-maker is good at approximating the destiny, this criterion, additionally called the optimism criterion, is used. It is a decision criterion taken by the decision maker who believes optimistically that nature will always be “on its side”. Maximax recommends maximizing the general payoff (Radzian, Itoh, & Inagaki, 2010). The constructive decision-maker considers the surest payoff for every potential direction of action. The restriction of those payoffs is detailed, and the subsequent order of action is selected. In these instances, they will result in horrific selections. A most rule consists of choosing the option that maximizes the maximum payoff. When the guaranteed payoff is slightly less than maximum possible payoff, and still the Maximax criterion is followed, it could be disastrous; shown below

	State of Nature			
Decision	1	2	MAX	
A	30	−10000	30	← maximum
B	29	29	29	

Maximin criterion: on the contrary, if the decision-maker is pessimistic at approximating the destiny, this criterion, also called the criterion of pessimism, is used. This is a more conservative approach to decision making where the nature is always assumed to be ”against us” regardless of the decision we make. Maximin manner that the minimal payoff is maximized. The terrible decision-maker situates the minimum payoff for every capability course of motion. The limit of those minimal payoffs is installed, and the following path of movement is selected. At each delivery stage, the investor will look at the worst possible outcome and pick out the highest. Therefore, the decision-maker selects the result guaranteed to mitigate his losses. He loses out on the danger of creating basic earnings inside the manner. Like in the below figure, alternative B would be selected using this rule but decision makers can prefer A as worst-case payoff is only slightly less than B.

	State of Nature			
Decision	1	2	MIN	
A	1000	28	28	
B	29	29	29	← maximum

Regret criterion: This criterion focuses on the regret of taking a selected direction of motion that the choice-maker may also have. This is based on the concept of opportunity loss. Regret is described as the difference between the first-class payoffs that we might have realized if we had acknowledged how the of nature would manifest and the payoff that became realized. This difference is also referred to as possible loss or opportunity value, which determines the volume of the loss due to no longer selecting the excellent alternative. The minimum regret approach is the one that minimizes the overall regret. The unforeseen conditions may occur when another option or an alternative appears in the payoff matrix and the decision can lead to inconsistent and incoherent results where the preferences can change or reverse.

3) Explain why - when using simulation - it is very important to provide summary statistics, such as the mean. In addition, explain the usefulness and approach of calculating confidence intervals.

4) Describe the purpose of sensitivity analysis in the context of Decision Theory in multistage decision making and how strategy tables and tornado charts can be used for this purpose.

5) Explain conceptually why the expected value of perfect information (EVPI) and the minimum expected opportunity loss (EOL) are equivalent.
The expected value of perfect information is the difference between the expected value with perfect information and highest expected value with existing information. It is simply the price paid to get perfect information. When a decision is taken, there could be uncertainty in it that it might go wrong.
EVPI = EPC – EMV; where EPC is expected payoff under certainty and EMV is expected monetary value
Whereas, minimum expected opportunity loss is the expected monetary value where the payoffs are substituted by regret values. Hence by definition, the expected opportunity loss always results in largest EMV and minimum expected opportunity loss result in expected monetary value under low regret payoffs. Hence
Minimum EOL = EPC – EMV
=>by the above two equations, EVPI = Minimum EOL.

References: Cliff Ragsdale, “Spreadsheet Modeling & Decision Analysis”, 6th Edition, Cengage Learning, 2015.
Hazra, A. (2017). Using the confidence interval confidently. Journal of thoracic disease, 9(10), 4125.
Radzian, A. M., Itoh, M., & Inagaki, T. (2010, November). Power transformer diagnosis decision-making based on unified combination rule and maximax-maximin criterion. In TENCON 2010-2010 IEEE Region 10 Conference (pp. 106-111). IEEE
Strong, M., Oakley, J. E., & Brennan, A. (2014). Estimating multiparameter partial expected value of perfect information from a probabilistic sensitivity analysis sample: a nonparametric regression approach. Medical Decision Making, 34(3), 311-326.

Part B – Application oriented questions (solve 3, 20 pts each)

1) An investment group is considering the construction of hotels around Funpark, a very popular amusement park. If demand is high, the investors estimate to make a profit of $250,000 per hotel. If demand is medium, they believe they would turn a profit of $105,000 per hotel; however, if demand is low, they would lose $80,000 per hotel. They believe that it is 25% certain that demand is high, and that there is a 30% chance it could be low. They are considering the construction of 3 or 2 hotels. Of course, they could also do nothing.
   Furthermore, the investment group would like to perform a market study. It would cost $25,000. The probability that the study results are negative is 0.3, that they are positive 0.7. Given a negative study, the probability for high demand is 0.1, for medium demand 0.4 and low demand 0.5. Given a positive study, the probability for high demand is 0.4, for medium demand 0.4 and for low demand 0.2.
   A) Using a decision tree, determine the best decision.
   B) The investment group realizes that the probabilities for the outcome of the study results are not certain. Perform a sensitivity analysis using a strategy table (vary the probability for the good outcome from 0.1 to 1). Discuss how the result of the sensitivity analysis changes your recommendation.

2) The investment group also wants to determine the optimal size of a hotel. The table below shows probabilities of generally reserved number of hotel rooms. Each room will be rented for $100. Statistics from the hospitality industry show that there is a 0.08 probability that guests that have booked a room close to an amusement park will not show up (and the hotel will lose the $100). On the other hand, there is also a 0.25 probability of last-minute reservations, i.e., the chance to rent out a room that was reserved but had a no-show. These rooms bring in $70. If a hotel is overbooked and a guest cannot get a room, $250 have to be paid to set up the guest in a different facility. Any room not sold results in $40 in variable costs. What is the optimal size of a hotel, in order to maximize daily profit? (Note: any other costs are ignored for this model).

Note: You can solve this either as an LP-Optimization problem or using PSISIMparam (one approach is sufficient).

Rooms reserved	50	51	52	53	54	55	56	57	58
Probability	0.05	0.15	0.18	0.18	0.19	0.1	0.08	0.04	0.03

3) Using the data in the file MSISFinalExamMovingServices, predict demand for the next three periods, using three different forecasting methods of your choice (moving average, weighted moving average with optimization of weights, exponential smoothing with optimization of alpha, simple regression). Your forecasts must include the determination of merit of the forecast and a discussion of your results.

Note: using for example a 2-period, 3-period and 4-period moving average does not account for three different models, but just one, namely moving average.

4) Dr. G. Ame, a psychologist in Las Vegas, is treating patients that are addicted to online gaming. The table below shows the number of patients he treated over the past 10 years, and the number of online games available.
   a. Develop a model to predict how many patients he will see next year. Develop a 95% confidence interval for the prediction. Discuss the merit of this prediction using a metric adequate to the model you used.
   b. The number of online games available is supposed to double each year. When he started, 9 online games existed. Based on this information, develop a model that predicts the patients he will next year. Develop a 95% confidence interval for the prediction. As above, discuss the merit of this model using an appropriate metric.
   c. Develop a model that combines both independent variables. Using an appropriate metric, discuss the merit of the model.
   d. Decide which model would be best for Dr. G. Ame to use.

Year	Number of patients	Number of online games
1	2	9
2	5	18
3	8	36
4	13	72
5	15	144
6	22	288
7	23	576
8	25	1152
9	28	2304

Answers:

1

A) The best decision is highlighted in the file 'Problem1_Adike.xlsx' where construction of 3 hotels is recommended in the case before making a market study and even after the investment group has decided to perform a market study, construction of 3 hotels is recommended in positive study and negative study.

Demand	Profit	Chance	3 Hotels	2 Hotels	Do Nothing
High	$250,000	25%	$750,000	$500,000	$0
Medium	$105,000	45%	$315,000	$210,000	$0
Low	($80,000)	30%	($240,000)	($160,000)	$0
		EMV	**257250**	**171500**	

The above table gives the EMV for the model without market study
This EMV is used to make analysis on the model with and without market study and the final EMV obtained from the decision tree is 2,63,900

B) A sensitivity analysis using a strategy table from the what if analysis in the solver and by varying the probability for the good outcome from 0.1 to 1 and the results are as below

Probability of good outcome	2
0.1	1
0.2	1
0.3	1
0.4	1
0.5	1
0.6	1
0.7	2
0.8	2
0.9	2
1	2

The final selection varies sensitively with the good outcome with the increase on probability from 70%. While the probability from 10% to 60 % is different.

3

The three different models were created in the file MSISFinalExamMovingSeries-1.xlsx' in three different sheets for moving average, weighted moving average and exponential smoothening.

Moving average method: The prediction for next three year using this method is as follows:

For 3 period moving average, y^22 = average (y22, y23, y24) = (y22+y23+y24)/3 = 108

y^23 = (y20+y21+y^22)/3 = 107.33 and y^24 = (y21+y^22+y^23)/3 = 110.11

Weighted moving average method: Let w1, w2, w3 be the weights for 3 period moving average which are the **decision variables** in minimizing MSE which is **objective function**. The sum of these weights must be 1 and the weights must lie between 0 and 1, which form as the **constraints** to the model. The solution is in the attached file. The prediction for next three years is as follows:

y^22 = y22*w3 + y23*w2 + y24*w1 = 107.899

y^23 = y20*w3 + y21*w2 + y^22*w1 = 111.050 and y^24 = y21*w3 + y^22*w2 + y^23*w1 = 109.652

Exponential smoothening: Let α be an exponent on which the forecast depends, signifying are the **decision variable** in minimizing MSE which is **objective function**. The decision variable must lie between 0 and 1, which form as the **constraints** to the model. The solution is in the attached file. The forecast for next three years is as follows

y^22 = y^21 + α(y21-y^21) = 108.69 and y for any year is to be 108.69 in this method, hence y^23 and y^24 = 108.69

The MSE for MA is 82.90, WMA is 70.36 and ES is 69.24
Since the error is small in Exponential smoothening, this method is more accurate than the other three.

4

A

To predict patients that doctor will see next year is attached in the excel file Problem4_Adike. From this model the regression equation generated is
Number of patients in an year = -1.25+3.3833*corresponding year. The R square and adjusted R square in this model are 98% and 97.8% respectively
-1.25 is the intercept and 3.38 is the slope coefficient in the model
The 95% confidence interval is developed in the sheet. The lower and upper limit is 28.80 and 36.36
The model is good as per the R2 square statistic which is the single most valuable factor and coefficient of determination. And the standard error is 1.37 which is low and good.

B

To predict patients that doctor will see next year based on the trend that online games are doubled upon every year is attached in the excel file Problem4_Adike. From this model the regression equation generated is
Number of patients in an year = 10.87+0.00937*Number of games online. The R square and adjusted R square in this model are 59% and 53.5% respectively
10.89 is the intercept and 0.00937 is the slope coefficient in the model
The 95% confidence interval is developed in the sheet. The lower and upper limit is 23 and 84 patients

C

The number of patients seen by the doctor based yearly increase of number of online games is predicted by the equation -2.36-0.0172*no. of online games+3.78*corresponding year
Where -2.36 is the negative intercept of the model with the slope on number of online games trending downwards in the graph (negative slope) and 3.78 slope coefficient of the year.
The R square for this model is 98.7% wit standard error of 98.3% which says that the model is well fitted.

D

The models generated in A and C parts are good as the R square is high with 98% approximately for both. Any of these two models would be better to use for the doctor to predict the number of patients.