SIT718 Real world Analytics

Trimester 1, 2018

PRACTICE EXAM PAPER 2

Special Instructions

This examination is CLOSED BOOK.
Calculators are ALLOWED.
Writing time is 2 HOURS.

Full marks can be obtained by correctly solving five (5) out of six (6) questions.
Formula sheet is attached at the end of the paper.

IMPORTANT NOTE:

- This is a PRACTICE EXAM PAPER - The aim is to give you an idea about the structure of the exam paper.
- The exam can include any part of the material. Programming with R and CPLEX is not required for the exam.

This examination question booklet must be handed in with any used answer booklets.

Question 1 *[20 marks] Aggregation functions - calculations*

(i) Answer the following questions about power means (PM). p is the parameter of power mean.

- If $PM(9, 10, 17, 16) = 13$ for some value of p, can we work out the value of $PM(12, 13, 20, 19)$ and $PM(18, 20, 34, 32)$ without knowing p?

(ii) Calculate the output for each of the following aggregation functions with respect to the input vectors, weighting vectors and additional input parameters indicated.

(a) $\text{WAM}(0.1, 0.9, 0.3, 0.2)$ with $\mathbf{w} = \langle 0.3, 0.6, 0.05, 0.05\rangle$.

(b) $\text{PM}_{\mathbf{w}}(0.7, 0.2, 0.5)$ with $\mathbf{w} = \langle 0.2, 0.25, 0.55\rangle$ and $p = -2$.

(c) $\text{GM}_{\mathbf{w}}(0.1, 0.2, 0.5, 0.2)$, with $\mathbf{w} = \langle 0.5, 0.3, 0.1, 0.1\rangle$.

(d) $\text{OWA}(0.2, 0.3, 0.1, 0.9)$, with $\mathbf{w} = \langle 7/16, 5/16, 3/16, 1/16\rangle$.

(iii) Consider the power mean $\text{PM}_{\mathbf{w}}$ with weighting vector $\mathbf{w} = \langle 0.1, 0.2, 0.1, 0.6\rangle$ and $p = 3$. For an input vector $\mathbf{x} = \langle 0.67, 0.5, 0.39, 0.2\rangle$, which out of x_2, x_4 should we increase if we want to improve the output as much as possible. Explain.

Question 2 *[20 marks] Interpreting parameters and weights*

[Note: Data Analysis is within the scope though this practice exam paper does not have a question on this.]

An online website uses customer preferences in order to evaluate the suitability of potential houses. Each house is rated according to the following criteria: Price (x_1), proximity to high achieving Schools (x_2) and proximity to public transport (x_3). All of the ratings are scored on a scale between 0 and 1, with 0 indicating not suitable at all and 1 indicating very suitable.

(a) Interpret the behaviour of the following aggregation functions with respect to their weighting vectors (which would correspond with preferences entered by the user) and parameters. Be sure to comment on the perceived importance of each criterion, whether the function favours high or low inputs, any redundancy between the variables.

i. A weighted arithmetic mean with $\mathbf{w} = \langle 0.1, 0.7, 0.2\rangle$.

ii. A weighted power mean with $p = -2$ and $\mathbf{w} = \langle 0.4, 0.4, 0.2\rangle$.

iii. An OWA function with $\mathbf{w} = \langle 0.6, 0.2, 0.2\rangle$.

iv. A Choquet integral with $v(\{1\}) = 0.8, v(\{2\}) = 0.6, v(\{3\}) = 0.6,$ $v(\{1, 2\}) = 0.8, v(\{1, 3\}) = 0.9, v(\{2, 3\}) = 0.6,$ $v(\{1, 2, 3\}) = 1$.

(b) Use *either* the power mean (given in ii) or the OWA (given in iii) to compare two houses (a and b) with ratings $\mathbf{x}_a = \langle 0.8, 0.2, 0.5\rangle$ and $\mathbf{x}_b = \langle 0.6, 0.5, 0.4\rangle$ and discuss the results.

Question 3 *[20 marks] Transformations*

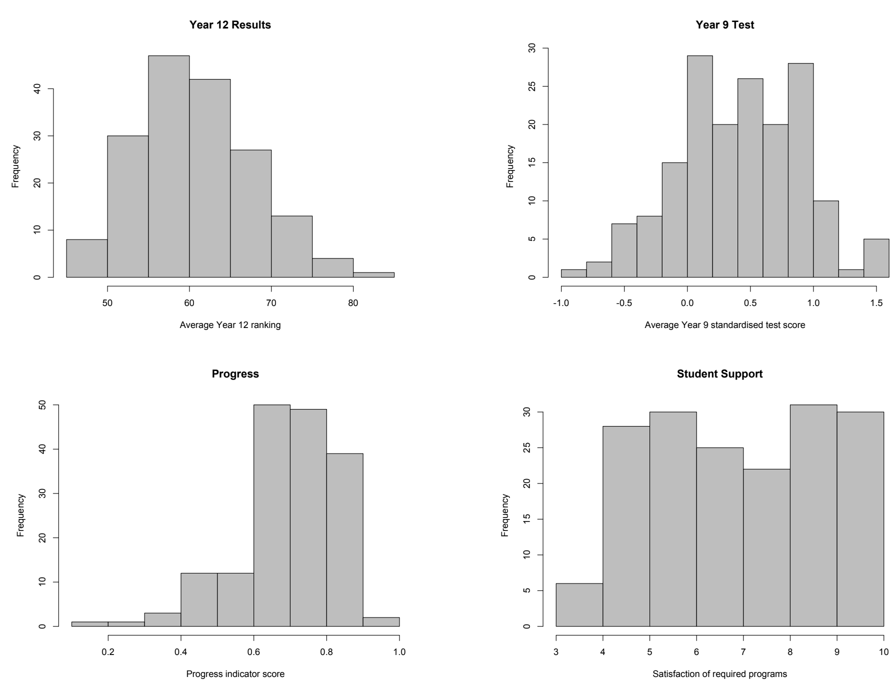
A total of 172 high schools across Melbourne are to be evaluated with respect to four criteria (with schools performing badly required to explain reasons for poor performance and an action plan for improvement).

Criterion 1: ***Year 12 final results*** - The average percentile ranking of the students who completed Year 12 the previous year. A score of 78.95 means that each student, on average, performed better than 78.95% of the state.

Criterion 2: ***Year 9 standardised Test results*** - The average score for Year 9 students at the school sitting a test (that is given to every student in Australia) in the current year. These scores are standardised so that if a school receives a score of 0, it means they are at the national average, whereas if they receive a score of -0.6, they are sitting -0.6 standard deviations below the national average.

Criterion 3: ***Student progress*** - This score is based on a number of available test scores. It is a score between 0 and 1 where 1 indicates that all students are progressing at a satisfactory rate, whereas 0 would indicate that no students are progressing. A score of 0.6 means that, overall, about 60% of students are progressing at the expected rate.

Criterion 4: ***Student support*** - This score is based on a 10 point checklist of programs and policies schools are required to adequately support disadvantaged students. A score of 10 indicates that all programs are in place and that no complaints have been made against the school. A score of 6 means that although some programs are in place, a number have not been implemented satisfactorily.



(a) For each variable, define appropriate transformations so that the score for each criterion ranges from 0 to 1 and the overall performance of each school can be evaluated using an aggregation function.

(b) Define an appropriate weighting vector that could be used and justify your choice of weights.

Question 4 *[20 marks] Linear Programming with more variables*

[Note: Linear Programming with 2 variables - graphical methods is within the scope though this practice exam paper does not have a question on this.]

Disunited Airlines will begin to operate in Australia starting from next year. They plan to employ ground agents in the most cost efficient way. There are 5 shifts in a day:

Shift 1: 06:00 - 14:00

Shift 2: 08:00 - 16:00

Shift 3: 12:00 - 20:00

Shift 4: 16:00 - 00:00

Shift 5: 10:00 - 06:00

The table below shows the hours covered by each shift, the various time periods for ground agents to operate, and the minimum number of ground agents required for each time period. The daily cost of hiring an agent for each time period is given in the last row of the table. Formulate an LP model to minimize the total cost of hiring of ground agents. [Hint: let x_j be a decision variable that determines the number of Shift j agents to hire, for $j = 1, \ldots, 5$.]

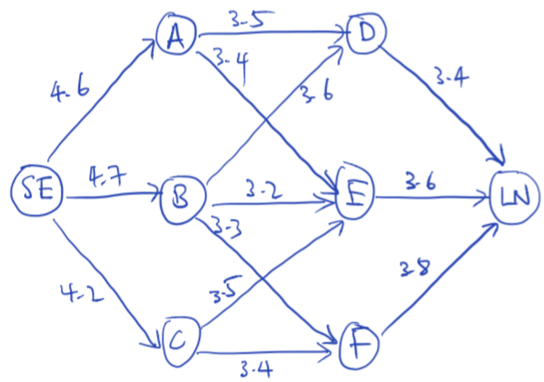
	Time periods covered					
Time Period	Shift 1	Shift 2	Shift 3	Shift 4	Shift 5	Min No. of agents needed
06:00 - 08:00	x					48
08:00 - 10:00	x	x				79
10:00 - 12:00	x	x				65
12:00 - 14:00	x	x	x			87
14:00 - 16:00		x	x			64
16:00 - 18:00			x	x		73
18:00 - 20:00			x	x		82
20:00 - 22:00				x		43
22:00 - 00:00				x	x	52
00:00 - 06:00					x	15
Daily cost per agent	\$170	\$160	\$175	\$180	\$195	

Question 5 *[20 marks] Game theory, Shortest Path, critical path Problems*

[Note: LP models of games with no saddle point are within the scope. Critical path problems are within the scope of the exam as well, even though this practice exam paper does not have a question on this.]

(i) Show that the following game has an optimal pure strategy pair. identify the optimal pure strategies for each of the players.

	A_1	A_2	A_3	A_4
a_1	5	2	1	3
a_2	3	3	-2	4
a_3	-1	3	-1	-3
a_4	2	-4	1	-2



(ii) One of Disunited Airlines' flights is about to take off from Seattle for a nonstop flight to London. There is some flexibility in choosing the precise route to be taken, depending upon weather conditions. The following network depicts the possible routes under consideration, where SE and LN are Seattle and London, respectively, and other nodes represent various intermediate locations.

The winds along each arc greatly affect the flying time (and so the fuel consumption). Based on current meteorological reports, the flying times (in hours) for this particular flight are shown next to the arcs. because the fuel consumed is so expensive, the management of Disunited Airlines has established a policy of choosing the route that minimizes the total flight time.

Use the Dijkstra's algorithm to find the path from Seattle to London with the shortest flight time.

Question 6 *[20 marks] Transportation Problems*

One of the main products of iHatePeas Company is canned peas. The peas are prepared at 3 canneries (near Goulburn Valley, Victoria; Mitchellstown, Whiteheads Creek, and Locksley), and then shipped to four major warehouses in Australia; Melbourne, Sydney, Brisbane, and Adelaide. As shipping costs are a major expense, management is initiating a study to reduce them as much as possible. The table below details the shipping costs, the supply from each cannery, and the demand of each city.

	Shipping costs ($) per Truckload				
Canneries	Melbourne	Sydney	Brisbane	Adelaide	Supply
Mitchellstown	464	513	654	867	75
Whiteheads Creek	352	416	690	791	125
Locksley	995	682	388	685	100
Demand	80	65	70	85	

Formulate an LP model to assign shipments to the various canneries–warehouses combinations that would minimize the total shipment cost.

[Note: Transportation problems with shortages or surplus are within the scope of the exam as well]

Formula Sheet:

Arithmetic Mean:

$$AM(\mathbf{x}) = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

Geometric Mean:

$$GM(\mathbf{x}) = \left(\prod_{i=1}^{n} x_i\right)^{1/n} = (x_1 x_2 \cdots x_n)^{1/n}.$$

Harmonic Mean:

$$HM(\mathbf{x}) = n\left(\sum_{i=1}^{n} \frac{1}{x_i}\right)^{-1} = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \cdots + \frac{1}{x_n}}.$$

Power Mean:

$$PM_p(\mathbf{x}) = \left(\frac{1}{n}\sum_{i=1}^{n} x_i^p\right)^{\frac{1}{p}} = \left(\frac{x_1^p + x_2^p + \cdots + x_n^p}{n}\right)^{\frac{1}{p}}.$$

Weighted Arithmetic Mean:

$$WAM_{\mathbf{w}}(\mathbf{x}) = \sum_{i=1}^{n} w_i x_i = w_1 x_1 + w_2 x_2 + \cdots + w_n x_n$$

Weighted Power Mean:

$$PM_{\mathbf{w}}(\mathbf{x}) = \left(\sum_{i=1}^{n} w_i x_i^p\right)^{\frac{1}{p}}$$

Weighted Geometric Mean:

$$GM_{\mathbf{w}}(\mathbf{x}) = \prod_{i=1}^{n} x_i^{w_i} = x_1^{w_1} \times x_2^{w_2} \times \cdots \times x_n^{w_n}$$

Weighted Harmonic Mean:

$$HM_{\mathbf{w}}(\mathbf{x}) = \left(\sum_{i=1}^{n} \frac{w_i}{x_i}\right)^{-1} = \frac{1}{\frac{w_1}{x_1} + \frac{w_2}{x_2} + \cdots + \frac{w_n}{x_n}}$$

Ordered Weighted Averaging (OWA):

$$OWA_{\mathbf{w}}(\mathbf{x}) = \sum_{i=1}^{n} w_i x_{(i)} = w_1 x_{(1)} + w_2 x_{(2)} + \cdots + w_n x_{(n)}$$

where $x_{(\cdot)}$ indicates that the arguments have been rearranged into non-decreasing order.

Orness of OWA=$\sum\limits_{i=1}^{n} w_i \frac{i-1}{n-1}$

Choquet integral (discrete):

$$C_v(\mathbf{x}) = \sum_{i=1}^{n} x_{[i]}(v(\{[i],[i+1],\ldots,[n]\}) - v(\{[i+1],[i+2],\ldots,[n]\}))$$

where $[i]$ means the inputs are arranged from lowest to highest and v are fuzzy measures.

Dijkstra's Algorithm:

Step 0

Set $= \emptyset$ (Initialise set of permanently labelled vertices)

Step 1

Set $P_1 = 0$ and $= \{1\}$ (Set permanent label of 1 to 0)
$\forall i \in V \setminus \{1\}$, s.t., $(1,i) \in A$, set $T_i = c_{1,i}$
$\forall i \in V \setminus \{1\}$, s.t., $(1,i) \notin A$, set $T_i = \infty$

Step 2

Repeat
Find $i^* = \arg\min_{i \in V \setminus} \{T_i\}$
Set $P_{i^*} = T_{i^*}$
Set $= \cup \{i^*\}$
$\forall i \in V \setminus$, s.t. $(i^*, i) \in A$, set $T_i = \min\{T_i, P_{i^*} + c_{i^*,i}\}$
Until $V \setminus = \emptyset$.

Transportation Problems–LP formulation:

$$\begin{aligned}
\min z = &(c_{11}x_{11} + c_{12}x_{12} + \cdots + c_{1n}x_{1n}) \\
&+ (c_{21}x_{21} + c_{22}x_{22} + \cdots + c_{2n}x_{2n}) \\
&+ \cdots + (c_{m1}x_{m1} + c_{m2}x_{m2} + \cdots + c_{mn}x_{mn}) \\
\text{s.t.} \quad & x_{i1} + x_{i2} + \cdots + x_{in} \leq \; s_i, \text{ one for each supply point } i \\
& x_{1j} + x_{2j} + \cdots + x_{mj} \geq \; d_j, \text{ one for each demand point } j \\
& x_{ij} \geq \; 0 \quad \text{for each } i = 1\ldots, m; \; j = 1, \ldots, n
\end{aligned}$$

where x_{ij} is the number of units shipped from supply point i to demand point j.