MAT9004 Assignment 1 (Answers)

1. A beekeeper estimates that the number of bees in a particular hive at the start of 2023 is 500, and that at the start of any subsequent day the number of bees in the hive has increased by 1% compared to what it was exactly one day earlier.

   (a) Find a function $b : \mathbb{R} \to (0, \infty)$ such that $b(t)$ models the population of the hive in the sense that $b(t)$ agrees with the beekeeper's estimates whenever $t \in \mathbb{N}$. Here $t \in \mathbb{R}$ measures the number of days since the start of the year (e.g. $t = 1/2$ is noon on New Year's day). Similarly, you should consider the population of bees to be a gradually growing real number (ignoring the fact that it is actually restricted to being an integer).

   ANS: Any time that a quantity grows (or shrinks) with a growth rate that is a constant proportion of the quantity, you are dealing with an exponential function. Thus we want $b(t) = Ar^t$ where r is the rate of growth (for our case $r = 1.01$) and the constant A determines the initial value (in our case $A = 500$). So the function we want is $b(t) = 500(1.01)^t$. [N.B. If you didn't know this about exponential growth, it would be easy to work out the pattern by working out the first few cases: $b(0) = 500$, $b(1) = 500 \times 1.01$, $b(2) = 500 \times 1.01 \times 1.01$ etc.] **[2]**

   (b) Explain why the inverse function b^{-1} exists.

   ANS: For a function to have an inverse it must be a bijection. There are two things to check, injectivity and surjectivity. To see that b is injective we check whether it sends two inputs, s and t to the same output. If $500(1.01)^t = 500(1.01)^s$ then $(1.01)^t = (1.01)^s$. Then taking log of both sides we get $t \log 1.01 = s \log 1.01$, which means that $t = s$. That means that there are NOT two different inputs that end up at the same output. In other words b is injective.

   To see that b is surjective we need to know that everything in the codomain is actually in the range. In this case, that means we need to know that for any given positive real number y we can find some real number x such that $b(x) = y$. We'll see this is the case in our answer to (c). **[2]**

   (c) Calculate a formula for b^{-1}.

   ANS: We solve $y = 500(1.01)^t$ for t. First we divide by 500 to get $y/500 = (1.01)^t$; then we take log of both sides, finding that $\log(y/500) = t \log(1.01)$ and hence $t = \log(y/500)/\log(1.01)$. We can thus take $b^{-1} : (0, \infty) \to \mathbb{R}$ to be defined by $b^{-1}(y) = \log(y/500)/\log(1.01)$. In particular, for any positive real number y we can calculate $x = \log(y/500)/\log(1.01) \in \mathbb{R}$ and it has the property that $b(x) = y$. **[2]**

   (d) The function b can be described as taking the time since the start of this year as input and outputting the population of bees in the hive at that time. Provide a similar description for what the function b^{-1} does.

   ANS: The function b^{-1} takes a hive population as input and outputs the time (measured in days from the start of 2023) at which the hive had that population. **[2]**

   (e) Find the derivative $b'(t)$. Explain what $b'(t)$ is measuring by giving a description of the same type that you gave in part (d).

   ANS: The derivative is $b'(t) = 500(1.01)^t \ln(1.01) \approx 47.655 \times (1.01)^t$. It takes a time (measured in days from the start of 2023) as input and outputs the rate of change in the population of the hive at that particular time. **[2]**

   (f) On which date does the population of bees in the hive reach 10000? Suppose that at the end of that day all but 1000 of the bees leave the hive to start a new colony elsewhere.

   ANS: The population reaches 10000 at the time $t = b^{-1}(10000) = \log(10000/500)/\log(1.01) \approx 301.07$. Note that this is on the 302-nd day of the year, since $t = 301$ is the end of the 301-st day. The 302 day is 29 October 2023, since $302 = 31 + 28 + 31 + 30 + 31 + 30 + 31 + 31 + 30 + 29$. **[2]**

(g) Write a function $c : [0, 365] \to (0, \infty)$ which gives the population of bees in the hive throughout 2023, given the migration described in (f). Assume that the population continues to increase at the same rate (1% per day) after the migration.

ANS: The original formula for $b(t)$ works when $t < 302$ and after that it is like we start again with 1000 bees (and with time now being counted since the start of the 302-nd day). Hence the function that we want is

$$c(t) = \begin{cases} 500(1.01)^t & t \in [0, 302), \\ 1000(1.01)^{t-302} & t \in [302, 365]. \end{cases}$$

[2]

(h) What is the population of bees in the hive at the end of the year?

ANS: Substituting $t = 365$ into the formula from (g) we get $1000(1.01)^{63} \approx 1871.7$. Rounding off we estimate the number of bees at the end of the year to be 1872. (1871 is satisfactory) **[2]**

2. READIT is a vendor of e-books. They provide recommendations to customers based on ratings by other customers. Each customer rates each book that they read, by giving a star rating from 1 star to 5 stars (there are no "half" stars). Suppose READIT wants to recommend books to a customer called Wendy. For each customer X other than Wendy, they first calculate

$$w(X) = \begin{cases} \frac{n}{10} & \text{if } n < 10, \\ 1 & \text{if } n \geqslant 10, \end{cases}$$

where n is the number of books that have been read (and rated) by both X and Wendy. Now for a book B that Wendy has not read they calculate a score $s(B)$ by the formula

$$s(B) = \frac{\sum\limits_{X \in C} w(X) r(X)}{\sum\limits_{X \in C} w(X)} \qquad (*)$$

where C is the set of customers that have read B, and $r(X)$ is the rating that X gave to B when they read it. READIT recommends the books with the highest score to Wendy.

(a) Under some circumstances it will not be possible to calculate the score $s(B)$ using the formula $(*)$. In order for $(*)$ to produce a real number as an answer, a particular set D of customers must be non-empty. What is D?

ANS: D is the set of customers who have read B and have also read at least one book that Wendy has read. If there exists a customer $X \in D$ then $w(X) > 0$ and $X \in C$ so the denominator in $(*)$ will be positive. If D is the empty set then the denominator will be zero, and dividing by zero will not produce a real number as a score. **[2]**

(b) Suppose that there is only one customer in the set D. What are the possible values of $s(B)$ in this situation?

ANS: If D only contains 1 customer, say T, then $w(X) = 0$ for every $X \in C$ except when $X = T$. That means that

$$s(B) = \frac{w(T)r(T)}{w(T)} = r(T).$$

Note that we are using that $w(T) \neq 0$ since $T \in D$. So the value of $s(B)$ equals the rating that T gave to B, which can only be one of $\{1, 2, 3, 4, 5\}$ (but could be any of these five values). **[2]**

(c) Assuming that D is not empty, what are the minimum possible value and the maximum possible value for $s(B)$?

ANS: Suppose that $w(X)$ is fixed for all customers $X \in C$. Then $(*)$ will increase (or stay unchanged) if $r(X)$ increases for any customer and decrease (or stay unchanged) if $r(X)$ decreases.

(The "unchanged option applies if $w(X) = 0$.) It follows that $s(B)$ is maximised by making all ratings $r(X) = 5$ and minimised by making all ratings $r(X) = 1$. If $r(X) = k$ for a constant k then

$$s(B) = \frac{\sum_{X \in C} w(X)k}{\sum_{X \in C} w(X)} = k \times \frac{\sum_{X \in C} w(X)}{\sum_{X \in C} w(X)} = k.$$

It is now obvious that the largest possible value of $s(B)$ is 5 and the smallest possible value is 1. **[2]**

(d) Explain why READIT uses the function $w(x)$ rather than just averaging the values of $r(X)$.

ANS: The purpose of $w(x)$ is to increase the "weighting" given to customers who have read more books that Wendy has read. If two people have read a lot of books in common then it probably indicates that they have similar taste in books and are more likely to like the same books. That makes it more likely that a recommendation from one will please the other. If two people have read very few books in common then a recommendation from one of these people to the other is not worthless, but it is less reliable than it would be if they had read a lot of the books in common. The function $w(X)$ ensures that less weighting is given to the ratings of books from customers whose tastes might not match Wendy's and more is given to those customers whose tastes probably do match Wendy's. **[2]**

3. The Laberal party are closely tracking their popularity through a 3 week election campaign. They find that on day d of the campaign the percentage of people who would vote for them is $v(d) = \frac{1}{200}(d^4 - 40d^3 + 594d^2 - 3888d + 13480)$. Consider v as a function of real values, $v : [0, 21] \to \mathbb{R}$.

(a) What are the domain and codomain of v?

ANS: The domain is $[0, 21]$ and the codomain is $\mathbb{R}$. **[1]**

(b) Find all stationary points of v.

ANS: We first need the derivative $v'(d) = \frac{1}{200}(4d^3 - 120d^2 + 1188d - 3888)$. Factoring this we find that $v'(d) = \frac{1}{50}(d - 9)^2(d - 12)$. Hence $v'(d) = 0$ when $d = 9$ and when $d = 12$ (and nowhere else). These are the two stationary points. **[2]**

(c) Classify each stationary point of v as a local minimum, local maximum or neither.

ANS: For this task we can use the second derivative $v''(d) = \frac{1}{50}(3d^2 - 60d + 297) = \frac{3}{50}(d - 9)(d - 11)$. At $d = 12$ we see that $v''(d) = 9/50 > 0$ and this ensures that $d = 12$ is a local minimum. On the other hand, we see that $v''(9) = 0$ which makes the second derivative test inconclusive. We thus need to try points close to $d = 9$ to check what the (first) derivative is doing. We see that $v'(8) = -2/25 < 0$ and $v'(10) = -1/25 < 0$ so the derivative has not changed sign at $d = 9$. It follows that $d = 9$ is neither a local minimum nor a local maximum. **[2]**

(d) What is the range of v?

ANS: To find the range we need to find the global minimum and global maximum. This requires checking the endpoints of the interval $[0, 21]$ as well as the stationary points. We see that $v(0) = 13480/200$, $v(9) = 4003/200$, $v(12) = 3976/200$ and $v(21) = 17827/200$. Since v is a smooth function, we infer that the range of v is $[3976/200, 17827/200]$. Put another way, the Laberal party's support varies from 19.88% to 89.14% (approx). **[2]**

(e) Over which intervals in its domain is v a convex function?

ANS: Here again we need the second derivative $v''(d) = \frac{3}{50}(d - 9)(d - 11)$. This is negative on $(9, 11)$ and positive on $[0, 9)$ and $(11, 21]$. In other words, $v(d)$ is convex for $d \in [0, 9)$ and $d \in (11, 21]$, and it is concave for $d \in (9, 11)$. N.B. For questions like this it is optional which endpoints you include in the intervals. **[2]**