

## Questions Q&A 6 CLV (Contractual) – Tuesday Dec 7<sup>th</sup>

### Questions Week 6 – CLV (Contractual)

- [Mieke];

- (1) Practice Quiz Q2: For this question, the discount rate is calculated using the following formula:  $d = 1.1^{(1/12)} - 1$

I do not understand how this formula is obtained. Also, it is given that we have a discount rate of 0.008, so why do we have to calculate this?

- (2) Assignment Q1: We are asked to provide the RLV immediately after renewal (the next payment is due the following month). Using slide 19, are we supposed to use the formula  $(m * p) / (1 + d - p)$  or  $(m * p) / (1 + d)^1$ ?

- (3) Assignment Q6: 'Up to how much would Blue Apron spend to rescue a customer who has renewed once, immediately after that first renewal? Use 200 time periods for your simulation'

The costs are given per week, while the data is given per month. Does this mean that we should use  $m = (\text{revenue} - \text{variable costs}) * 4$ ?

Also, it is not totally clear to me what tau represents. Could you please clarify this? For this case, I assumed it to be  $\tau = 1$ . Is this correct?

- [Noa];

- (1) When we determine t as a sequence, we sometimes start at 0 ( $t \leftarrow \text{seq}(0, 10)$ ) and sometimes start at 1 ( $t \leftarrow \text{seq}(1, 200)$ ). When should we start at 0 and when at 1?

*See Lab Session just above **Fit** where we use  $t \leftarrow \text{seq}(0, 10)$  for 10 periods and Lab Session **at Calculating CLV with SBG** where we use  $t \leftarrow \text{seq}(1, 200)$ .*

For the former we argue there are 10 periods, whereas (0,10) gives 11 'time-points', right? I'm confused when to start at 0 and when at 1 + how should we determine the number of time periods? (e.g. 10 time periods is from 0 till 10 but that gives 11 time periods?)

- (2) Practice Quiz - Question 4 and 6:

In Q4 we use  $S\_pred \leftarrow c(1, \text{cumprod}(r\_pred)[1:199])$  and in Q6 we use  $S\_pred \leftarrow \text{cumprod}(r\_pred)$ . When do we use which?

- (3) Practice Quiz - Question 2:

To get a and b in the sBG we use

`lost <- -diff(active_cust[1:5])`

`active <- active_cust[2:5]`

What is the reason for using 1:5 for lost and using 2:5 for active?

- [Jorg];
  - (1) In the Q&A after the computerlab we discussed the following comprehension check:  
Comprehension check 4 (Webclip 6.4): *"Add the actual retention rate and that based on the geometric model. How well do these models fit the data? Add the survival function and make the same comparison."*

However, we only discussed how to do predict the sBG for the retention rate. Could you show how you can also predict the sBG this for the survivor rate?