How User Ratings Change with Time: Theoretical Explanation of an Empirical Formula

Julio Urenda, Manuel Hernandez,
Natalia Villanueva-Rosales, and Vladik Kreinovich

University of Texas at El Paso
El Paso, TX 79968, USA
jcurenda@utep.edu, mahernandez46@miners.utep.edu
nvillanuevarosales@utep.edu, vladik@utep.edu

1. Formulation of the Problem

- For many industries, it is important to predict the user's reaction to new products.
- To make this prediction, it is reasonable to use:
  - the past records of the user's degree of satisfaction
  - with different similar products.
- One of the problems with such a prediction is that the user's degree of satisfaction may change with time.
- The first time you see an exciting movie or read an exciting book, you feel very happy.
- When you see this movie the second time: you may notice:
  - holes in the plot or
  - outdated (and thus, somewhat clumsy) computer simulation.

2. Formulation of the Problem (cont-d)

- Because of this phenomenon, for each user, the ratings of the same product decrease with time.
- In other words:
  - instead of the formula $r = r(u, p)$ describing how rating depends on the user u and the product p,
  - a more accurate estimates can be obtained if we take this dependence into account:

$$r = r(u, p) + c_u(t).$$

- Here a decreasing function $c_u(t)$ – which is, in general, different from different users – change with time.
- To test different ratings models, in the 2000s, Netflix had a competition.

3. Formulation of the Problem (cont-d)

- The winning model used an empirical formula

$$c_u(t) = \alpha_u \cdot \mathrm{sign}(t - t_u) \cdot |t - t_u|^{\beta_u}.$$

- Here t_u is the mean date of rating, and α_u and β_u are parameters depending on the user.
- Actually, it turned out that the value β_u is approximately the same for all the users.
- The question is: why this formula works well, while possible other dependencies do not work so well?

4. First Technical Comment

- The above formula does not uniquely determine the functions $r(u,p)$ and $c_u(t)$:
  - we can add a constant to all the ratings $r(u,p)$ and subtract the same constant from all the values $c_u(t)$
  - and still get the same overall ratings r.
- To avoid this non-uniqueness, we can, e.g., select $c_u(t)$ in such a way that $c_u(t_u)=0$.
- This is, by the way, exactly what is done.
- This equality is easy to achieve:
  - if we have a function $c_u(t)$ for which $c_u(t_u)\ne 0$,
  - then we can consider new functions

$$\widetilde{c}_u(t_u) \stackrel{\text{def}}{=} c_u(t)-c_u(t_u) \text{ and } \widetilde{r}(u,p) \stackrel{\text{def}}{=} r(u,p)+c_u(t_u).$$

5. First Technical Comment (cont-d)

- Then, as one can easily see, we have

$$\widetilde{r}(u,p)+\widetilde{c}_u(t)=r(u,p)+c_u(t).$$

- So, all the predicted ratings remain the same.
- In view of this possibility, in the following text, we will assume that $c_u(t_u)=0$.

6. First Idea: The Description Should Not Depend on the Unit for Measuring Time

- We are interested in finding out how the change in ratings depends on time.
- In a computer model, time is represented by a number.
- However, the numerical value of time depends on what starting point we choose and what unit we use for measuring time.
- In our situation, there is a fixed moment t_u.
- So it is reasonable to use t_u as the starting point and thus use $T \stackrel{\text{def}}{=} t - t_u$ to measure time instead of t.
- In the new scale, the formula describing how ratings change with time takes the form $C_u(T)$, so that
$$c_u(t) = C_u(t - t_u).$$
- The condition $c_u(t_u) = 0$ takes the form $C_u(0) = 0$.

7. First Idea (cont-d)

- In terms of the new time scale, the empirically best formula leads to the following expression for $C_u(T)$:

$$C_u(T) = \alpha_u \cdot \mathrm{sign}(T) \cdot |T|^{\beta_u}.$$

- While there is a reasonable starting point for measuring time, there is no fixed unit of time.
- We could use years, months, weeks, days, whatever units make sense. If:
  - we replace the original measuring unit with a new unit which is λ times smaller,

$$T \to \widetilde{T} = \lambda \cdot T.$$

- There is no special measuring unit for time.
- So, it makes sense to require that the formulas not change when we make a different selection.

8. This Has to Be Related to a Change in How We Measure Ratings

- Of course, we cannot simply require that the formula for $C_u(T)$ be invariant under the change $T \to \lambda \cdot T$.
- Indeed, in this case, we would have $C_u(\lambda \cdot T) = C_u(T)$ for all λ and T.
- Thus, for each $T_0 > 0$, by taking $T = 1$ and $\lambda = T_0$, we would be able to conclude that $C_u(T_0) = C_u(1)$.
- So, instead of the desired decreasing function, we would have a constant function $C_u(T) = \text{const}$.
- This seeming problem can be easily explained if we recall how scale-invariance works in physics.

9. How We Measure Ratings (cont-d)

- The formula $v = d/t$ describing how the velocity v depends on the distance d and time t:
  - clearly does not depend
  - on what measuring unit we use for measuring distance.
- We could use meters, we could use centimeters, we could use inches.
- However, this does not mean that:
  - if simply change the measuring unit for distance and thus replace d with $\lambda \cdot d$,
  - the formula remains the same.

10. How We Measure Ratings (cont-d)

- For the formula to remain valid, we also need to accordingly change the unit for measuring velocity, e.g.
  - from meters per second
  - to centimeters per second.
- Similarly in our case, when we describe the dependence of rating on time:
  - we cannot just change the unit for time,
  - we also need to change another unit – which, in this case, is the unit for ratings.
- But does this change make sense?
- At first glance, it may seem that it does not.
- We ask the user to mark the quality of a product (e.g., of a movie) on a certain fixed scale (e.g., 0 to 5).

11. How We Measure Ratings (cont-d)

- So how can we change this scale?
- Actually, we can.
- Users are different.
- Some users use all the scale, and mark the worst of the movies by 0, and the best by 5.
- What happens when a new movie comes which is much better than anything that the user has been before?
- In this case:
  - the user has no choice but to give a 5 to this movie as well,
  - wishing that the scale had 6 or 7 or even more.

12. How We Measure Ratings (cont-d)

- Similarly:
  - if a movie has a very negative experience with a movie,
  - a much worse one than anything that he or she has seen before,
  - this user places 0 and wishes that there was a possibility to give -1 or even -2.
- Other users recognize this problem and thus:
  - use only, e.g., grades from 1 to 4,
  - reserving 0 and 5 for future very bad and very good products.

13. How We Measure Ratings (cont-d)

- Some professors grade the student papers the same way:
  - using, e.g., only values up to 70 or 80 out of 100, and
  - leaving higher grades for possible future geniuses.
- In other words:
  - while the general scale – from 0 to 5 or from 0 to 100 – is indeed fixed,
  - the way we use it changes from one user to another one.
- Some users use the whole scale, some "shrink" their ratings to fit into a smaller sub-scale.
- A natural way to describe this shrinking is by an appropriate linear transformations.

14. How We Measure Ratings (cont-d)

- This is how, e.g., we estimate the grade of a student who had to skip a test worth 20 points out of 100:
  - if overall, the student earned 72 points out of 80,
  - we mark it as $\dfrac{72}{80} \cdot 100 = 90$ points on a 0 to 100 scale.
- Depending on what scale we use for ratings, the corresponding rating values change by a linear formula:

$$r \to \widetilde{r} = a \cdot r + b.$$

- For the difference between the ratings, we get:

$$r_1 - r_2 \to (a \cdot r_1 + b) - (a \cdot r_2 + b) = a \cdot (r_1 - r_2).$$

15. How We Measure Ratings (cont-d)

- So:
  - when we change the unit for measuring time by a λ times smaller one,
  - we may need to according re-scale our difference $C(T)$ between the ratings.
- Thus, we arrive at the following precise formulation of the desired invariance.

16. Formal Description of Unit-Independence

- We want to select a function $C_u(T)$ for which:
  - for each $\lambda > 0$,
  - there exists a value $a(\lambda)$ for which

$$C_u(\lambda \cdot T) = a(\lambda) \cdot C_u(T).$$

- It is also reasonable to assume that:
  - the function $C_u(T)$ continuously change with time,
  - or at least change with time in a measurable way.

17. What Can We Conclude Based on This Independence

- It is known that:
  - every measurable (in particular, every continuous) solution to the equation
  $$C_u(\lambda \cdot T) = a(\lambda) \cdot C_u(T) \text{ for } T > 0$$
  - has the form
  $$C_u(T) = \alpha_u^+ \cdot T^{\beta_u^+} \text{ for some } \alpha_u^+ \text{ and } \beta_u^+.$$
- Similarly, for $T < 0$, we get
$$C_u(T) = \alpha_u^- \cdot T^{\beta_u^-} \text{ for some } \alpha_u^- \text{ and } \beta_u^-.$$
- These formulas are similar to the desired formula.
- However, we still have too many parameters: four instead of the desired two.
- To get the desired formula, we need one more idea.

18. Second Idea: The Change in Rating Should Be the Same Before And After t_u

- It is reasonable to require that:
  - for each time interval $T > 0$,
  - the change of rating should be the same before and after t_u.
- In other words:
  - the change of ratings between the moments $t_u - T$ and t_u should be the same as
  - the change of ratings between the moments t_u and $t_u + T$.
- The change of ratings between the moments $t_u - T$ and t_u is equal to

$$c_u(t_u) - c_u(t_u - T) = -(c_u(t_u - T) - c_u(T)) = -C_u(-T).$$

19. Second Idea (cont-d)

- The change of ratings between the moments $t_u + T$ and t_u is $c_u(t_u + T) - c_u(t_u) = C_u(T)$.
- Thus, the above requirement means that for every $T > 0$, we should have $-C_u(-T) = C_u(T)$.
- Thus, for all $T > 0$: $\alpha_u^+ \cdot T^{\beta_u^+} = -\alpha_u^- \cdot T^{\beta_u^-}$.
- Since this must be true for all T, we must have

$$\alpha_u^+ = -\alpha_u^- \text{ and } \beta_u^+ = \beta_u^-.$$

- Thus, for both $T > 0$ and $T < 0$, we indeed have $c_u(T) = \alpha_u \cdot \mathrm{sign}(T) \cdot |T|^{\beta_u}$, with $\alpha_u = \alpha_u^+$ and $\beta_u = \beta_u^+$.
- The empirical formula is explained.

20. Acknowledgments

This work was supported in part by the National Science Foundation grants:

- 1623190 (A Model of Change for Preparing a New Generation for Professional Practice in Computer Science),
- HRD-1242122 (Cyber-ShARE Center of Excellence).