Preregistration

Epidemiology Beyond Epidemiology

Prof. Lucas Helal, MMSc, PhD[1]

[1] Graduate Program in Epidemiology, Universidade Federal do Rio Grande do Sul

04. September 2023

Why Epidemiology Beyond Epidemiology?

Let's use an example... The SIR Model

The Exponential Growth of An Epidemics

The spread of an epidemic disease depends on both the amount of contact between individuals and also on the possibility of an infected person transmitting the disease to another person. This is can be represented by a geometric series in with **the ratio of each two consecutive terms is a constant** r. The n-th term is as follows:

$$A(n) = A(0) \cdot r^n \tag{0.1}$$

Consider the initial value to be $A(0)$. For any $r > 1 \ \wedge \ A(0) > 0 \ \therefore \ A(n) \overset{+++}{\Longrightarrow} +\infty$ - that means, goes towards the infinity quickly. It $0 < r < 1$, then $A(n) \overset{---}{\Longrightarrow} 0$. So, there is the exponential function:

$$x = r^t \tag{0.2}$$

Differentiation of the exponential function

$$\frac{dr}{dt} = \lim_{h \to 0} \frac{r^{t=h} - r^t}{h} = r\hat{\ }t \lim_{h \to 0} \frac{r^h - 1}{h} \tag{0.3}$$

The Naipier's constant e satisfies this equation. Then:

$$\frac{d}{dt} \cdot e^t = e^t \ x(t) = C \cdot e^{at} \tag{0.4}$$

The inverse function can represented by $t = log_r \cdot x$.

> Let's verify how the spread behaves as the time increases, by using t = 2.

$$\frac{x(t_2)}{x(t_1)} = \frac{e^{at_2}}{e^{at_1}} = e^{at_2 - t_1)} = 2 \tag{0.6}$$

And $a(t_2 - t_1) = ln \cdot 2$.

Let's check a linear and semi-log plots to better visualize the behavior of the function with a two-fold increasing in time.

"'{'text}

library(ggplot2) x <- seq(0,10,1) y <- 2.71^x xy <- as.data.frame(cbind(x,y)) pxy <- plot(x,y)

ggplot(xy, aes(x=x,y=y)) + geom_point() + geom_smooth(span = 0.4, se=FALSE) + xlab("Time") + ylab("Infected Cases") "'

The SIR Model was originally proposed by Kermack and McKendrinc in 1927 (1). It is a compartmental model that assumes a given population N during an epidemic may be reasonable divided into three sub-populations: S for susceptible; I for infected; and R for recovered,

We may interpret S as the population at risk to be infected. The compartmental model may be pictured as follows:

Figure 1. Compartmental model of SIR.

Compartmental model of SIR

A. $S = S(t)$ is a function of S_i - S susceptible individuals, given the time t.

$$N_S \sim S(t) \tag{1.0}$$

$$N_I \sim I(t) \tag{1.1}$$

$$N_R \sim R(t) \tag{1.2}$$

$$N_{total} \sim S(t) + I(t) + R(t) \tag{1.3}$$

Then, we do have our dependent variables, which are a share of our individual's populations. For a given $N_{S,I,R}$, there is a fraction of each N assuming N as the true populational prevalence - parameter $\hat{P}$. This is denoted by:

$$\frac{S(t) \ \vee \ I(t) \ \vee \ R(t)}{N}, \ \exists \ \ s(t) \ \vee \ i(t) \ \vee \ r(t) \tag{1.4}$$

Epidemiology and Rates

This is a fundamental problem about **rates**. Essentially, during an epidemic episode, we are not too much interest in what function f fits better the observations, but, the **behavior** of the phenomena, This is crucial to understand incidences, prevalence, probability distributions or, even though, forecasts. Nonetheless, as any real-world event that is approximated by a numerical expression, it is susceptible to human/non-human countermeasures which **frequently overcomes** the theoretical modelling of a phenomena.

Bringing rates to Epidemiology

Numbers has an unique property to permit us an abstraction level to solve problems that fit in a plethora of ordinary real-world scenarios. For example, let's take the following situations, assuming we have success for all of them (i.e., the expected outcome):

- OK, but at which manner these bacteria grow in my Petri plaque over time?
- The inflation is out of control in my country over the year. Can I understand how is it going on?
- A professor of mine told me in an Epi class that the world is going better. As far as I remember, fertility rates, maternal and child death rates, income distribution (Gini's index and Lorenz curves) and inequity reduction through poverty, gender and race discrimination or hunger are good candidates to proxy an evolution. Can I estimate a magnitude of this potential shift globally?
- Frequently, I am working with discrete variables/integers. However, sometimes it is pragmatically easier for me to work with some data in a continuous manner. My Stats' course regent told me to transform it. But to transform to what? To transform how? Digging in the web, I've came across integral transformations, like the linear Laplace's transformation. There is indeed a fancy L to denote it ($\mathcal{L}$). Curiously, may professor's name is Lucas and begins with the L. How does it works?

All of these problems arrives or make use of the concept of rates. Using our abstraction feature, some mathematical properties not so difficult to calculate are numerically (and sometimes behaviorally) equal to a rate of a function f in a point $\overline{a}$.

I'm not going to dig in the demonstration of from where a rate of a function f in a point $\overline{a}$ can be represented. I would like to stress some core ideas:

- First, imagine any curve plotted in a Cartesian coordinate system
- Second, please note the proportion between a value y_i to a given x_i observation is not constant. For each unit of x_n increased or decreased, its dependent variable y_n - or simply $f(x)$, does not move the very same one unit then. Sometimes the response in y is greater (*"speeds the function"*); sometimes is lower (*"slows the function"*)
- Third, I know it's not sufficient to understand rates of a function. Moreover, you probably is asking yourself how can you calculate a rate in a **given point** if your learned in high school (correctly) that rates are differences between two moments, broadly speaking (**then, between at least two points**)?
- This is when the concept of limit of a function towards a point b gains a lot of importance to understand rates and how we can calculate them in a **given point** of a function (**yes!**). You need to use the abstraction now.

Understanding what is a rate of a function in a given point and how to calculate it

First, let's comment about limits.

A limit of a function $f(x)$ when $x = a$ and will move to a position in $x\ != $ then a, we are trying to understand the **expected value** of $f(x)$ when arriving in the next point - **if arrives...** . So, limit is more a concept and an idea than a mathematical feature to solve problems. Keep it mind and I promise your life will be way easier.

Formally, the limit of a function $f(x)$, arriving in a point b, is denoted by:

$$\lim_{x \to\ b} f(x) = y^{"} \tag{1.5}$$

This is:

Limit of y when x tends to b is equal to y"

I know I'm jumping several steps, but now, instead of a given curve, please imagine a circle in Cartesian coordinate system. Any circle.

- You learned that a tangent line is a line that touches a given curve in a single point. It can be true, but it also cannot. Indeed, this concept for a tangent line in a given point is incomplete and is an heuristic. It works for a restricted working space, and for circles/circumferences. But, if you realize that lines (or curves) goes towards infinit (unless interrupted by any rule), it's fair enough to picture this line may touch another point in a place over the infinit. Why this concept is important?

- Because a tangent line is:

  A tangent line of a given point of a curve is the line that best matches with the curvature at the tangency point.

  Indeed, in a tangency point, we fairly approximate the shape of the curve as (very nearly) equal to the shape of the line.

Therefore, the line and the curve at that point can be considered the same thing. And the major implication of this:

> The properties of the original line can be extrapolated to the curve in the tangency point.

Indeed, quoting my professor and hard sciences mentor, Prof. Julio Lombaldo, if you zoom-in the tangent point *ad infinitum*, you will realize that the curvature in that given point will be approximately parallel and, even more, narrowed to the original straight line.

From Limits to Rates

So, now I acknowledge you are ready to understand broadly what is and how to calculate a rate of a function in a given point.

Now, let's work with two points in a circumference, instead of one. If a straight line touches a circumference in at least two points simultaneously, we are dealing with you learned as a secant straight line.

Thus, with two points, you can calculate a rate of a function at that segment of line, correct? You will compare their positions at the y axis to the x axis - like time-dependent functions of first order we learned in Physics during the high school:

$$V_m = \frac{\Delta S}{\Delta t} \tag{1.6}$$

I bet it seems familiar to you. The equation to estimate the mean velocity of given two different positions and the time spent to move from the initial to the final one.

Well, that's a rate. And rates like this are fairly easy to calculate. However, in the vast majority of the time in Epidemiology, we are interest in the rate at that given moment, not the mean rate; i.e., the estimate in a single point, not between two of them; i.e., again, the instantaneous rate, when $t = c$, or, formally speaking, the time-window between the moments of interest tends to zero: $\Delta t \to 0$. There comes $\lim_{x \to c} f(x)$ again.

Indeed, if you realize you can approximate the intended position close or equal to x, you will have two points "virtually" at the same position and will be able to calculate a rate in a single point. To this, we name the property as the **derivative** of a given function $f(x)$, commonly represent by $\frac{\delta x}{\delta}$; $f'(x)$; $\frac{dx}{dt}$ or, more infrequently, $\dot{y}$.

We do have two classical manners to demonstrate what a derivative of a function is:

- by limits of a function;
- by analytically geometry.

Importantly, both manners adequately represent philosophically the majority of the phenomena in Epidemiology, and that's what we need to keep in mind.

«««< DEMONSTRACAO POR LIMITES »»»» «««< DEMONSTRACAO POR GEOMETRIA »»»

«««< EXEMPLOS DA MATEMATICA SUPERIOR EM EPIDEMIOLOGIA »»»