Why Do We Need Two Doses of Covid-19 Vaccine: A Qualitative Explanation

Laxman Bokati[1], Julio Urenda[2,3], Olga Kosheleva[4], and Vladik Kreinovich[1,3]

[1]Computational Science Program
[2]Department of Mathematical Sciences
[3]Department of Computer Science
[4] Department of Teacher Education
University of Texas at El Paso
500 W. University
El Paso, TX 79968, USA
lbokati@miners.utep.edu, jcurenda@utep.edu,
olgak@utep.edu, vladik@utep.edu

Abstract

It is known that the most effective protection from Covid-19 comes if the vaccination is done in two doses separated by several weeks. In this paper, we provide a qualitative explanation for this empirical fact.

Mathematics Subject Classification: 92C60 60F05 58D19

Keywords: Covid-19 vaccination, scale-invariance, shift-invariance, Central Limit Theorem

1 Formulation of the Problem

For Covid-19, the most efficient way is not to get one large doze, but to get two doses separated by a few weeks. To be more precise, one dose may be good for a short-term effect, but two doses are needed to maintain a long-term effect.

But why? In this paper, we provide a qualitative explanation for this empirical fact.

Comment. Interestingly, the corresponding mathematics turns out to be similar to the one presented in [1] for a completely different case – of short-term vs. long-term effects of different learning strategies.

2 Analysis of the Problem and the Resulting Explanation

Main idea. According to evolution theory, nature wants to maximize the expected gain – i.e., equivalently, to minimize the expected loss.

Notations. To describe this idea in precise terms, let us introduce some notations.

- Let m denote the efforts needed to store and maintain antibodies.
- Let L denote the expected loss when the virus comes and there are no antibodies ready to fight it.
- Finally, let p denote our estimate of the probability that the virus will appear during the given time interval (short-term or long-term).

If the body stores and maintains the antibodies, we have loss m. If it does not, then the expected loss is equal to $p \cdot L$. So, the body will store and maintain antibodies if the second loss of larger, i.e., if $p \cdot L > m$, i.e., equivalently, if

$$p > \frac{m}{L}.$$

Discussion. To compare two different vaccination strategies, we need to compare the corresponding probability estimates p.

Let us formulate the problem of estimating the corresponding probability p in precise terms.

Towards a precise formulation of the probability estimation problem. In the absence of other information, to estimate the probability that the antibodies will be needed in the future, the only information that the body can use is that there were two moments of time at which we needed these antibodies were needed in the past:

- the moment t_1 when the body gets the first dose of the vaccine, and
- the moment t_2 when the body gets the second dose.

If we get both doses in a single shot, then the moment t_2 is close to t_1, so the difference $t_2 - t_1$ is small. If the second dose is delayed, the difference $t_2 - t_1$ is larger.

Based on this information, the body has to estimate the probability that there will be another moment of time during some future time interval when antibodies will be needed. How can it do that?

Let us first consider a deterministic version of this problem. Before we start solving the actual probability-related problem, let us consider the following simplified deterministic version of this problem:

- the body knows the times $t_1 < t_2$ when the antibodies were needed;
- the body needs to predict the next time t_3 when the antibodies will be needed.

We can reformulate this problem in more general terms:

- an agent (in this case, the body) encountered some event at moments t_1 and
$$t_2 > t_1;$$
- based on this information, it wants to predict the moment t_3 at which the same event will happen again.

In other words, we need to have a function $t_3 = F(t_1, t_2) > t_2$ that produces the desired estimate.

What are the reasonable properties of this prediction function? The numerical value of the moment of time depends on what unit we use to measure time – e.g., hours, days, or months. It also depends on what starting point we choose for measuring time. We can measures it from Year 0 or – following Muslim or Buddhist calendars – from some other date.

If we replace the original measuring unit with the new one which is a times smaller, then all numerical values will multiply by a: $t \to t' = a \cdot t$. For example, if we replace seconds with milliseconds, all numerical values will multiply by 1000, so, e.g., 2 sec will become 2000 msec. Similarly, if we replace the original starting point with the new one which is b units earlier, then the value b will be added to all numerical values: $t \to t' = t + b$. It is reasonable to require that the resulting prediction t_3 not depend on the choice of the unit and on the choice of the starting point. Thus, we arrive at the following definitions.

Definition 1. *We say that a function $F(t_1, t_2)$ is* scale-invariant *if for every t_1, t_2, t_3, and $a > 0$, if $t_3 = F(t_1, t_2)$, then for $t'_i = a \cdot t_i$, we get $t'_2 = F(t'_1, t'_2)$.*

Definition 2. *We say that a function $F(t_1, t_2)$ is* shift-invariant *if for every t_1, t_2, t_3, and b, if $t_3 = F(t_1, t_2)$, then for $t'_i = t_i + b$, we get $t'_2 = F(t'_1, t'_2)$.*

Proposition 1. *A function $F(t_1,t_2) > t_2$ is scale- and shift-invariant if and only if it has the form $F(t_1,t_2) = t_2 + \alpha \cdot (t_2 - t_1)$ for some $\alpha > 0$.*

Proof. Let us denote $\alpha \stackrel{\text{def}}{=} F(-1,0)$. Since $F(t_1,t_2) > t_2$, we have $\alpha > 0$. Let $t_1 < t_2$, then, due to scale-invariance with $a = t_2 - t_1 > 0$, the equality $F(-1,0) = \alpha$ implies that $F(t_1 - t_2, 0) = \alpha \cdot (t_2 - t_1)$. Now, shift-invariance with $b = t_2$ implies that $F(t_1,t_2) = t_2 + \alpha \cdot (t_2 - t_1)$.

The proposition is proven.

Discussion. Many physical processes are reversible: if we have a sequence of three events occurring at moments $t_1 < t_2 < t_3$, then we can also have a sequence of events at times $-t_3 < -t_2 < -t_1$. It is therefore reasonable to require that:

- if the agent's prediction works for the first sequence, i.e., if, based on t_1 and t_2, we predict t_3,
- then the agent's prediction should work for the second sequence as well, i.e. based on $-t_3$ and $-t_2$, we should predict the moment $-t_1$.

Let us describe this requirement in precise terms.

Definition 3. *We say that a function $F(t_1,t_2)$ is* reversible *if for every t_1, t_2. and t_3, the equality $F(t_1,t_2) = t_3$ implies that $F(-t_3,-t_2) = -t_1$.*

Proposition 2. *The only scale- and shift-invariant reversible function $F(t_1,t_2)$ is the function $F(t_1,t_2) = t_2 + (t_2 - t_1)$.*

Comment. In other words, if the agent encounters two events separated by the time interval $t_2 - t_1$, then the natural prediction is that the next such event will happen after exactly the same time interval.

Proof of Proposition 2. In view of Proposition 1, all we need to do is to show that for a reversible function we have $\alpha = 1$. Indeed, for $t_1 = -1$ and $t_2 = 0$, we get $t_3 = \alpha$. Then, due Proposition 1, we have

$$F(-t_3,-t_2) = F(-\alpha,0) = 0 + \alpha \cdot (0 - (-\alpha)) = \alpha^2.$$

The requirement that this value should be equal to $-t_1 = 1$ implies that $\alpha^2 = 1$, i.e., due to the fact that $\alpha > 0$, that $\alpha = 1$.

The proposition is proven.

From simplified deterministic case to the desired probabilistic case. In practice, we cannot predict the actual time t_3 of the next occurrence, we can only predict the *probability* of different times t_3.

Usually, the corresponding uncertainty is caused by a joint effect of many different independent factors. It is known that in such situations, the resulting probability distribution is close to Gaussian – this is the essence of the Central

Limit Theorem which explains the ubiquity of Gaussian distributions; see, e.g., [2]. It is therefore reasonable to conclude that the distribution for t_3 is Gaussian, with some mean μ and standard deviation σ.

There is a minor problem with this conclusion; namely:

- Gaussian distribution has non-zero probability density for all possible real values, while
- we want to have only values $t_3 > t_2$.

This can be taken into account if we recall that in practice, values outside a certain $k\sigma$-interval $[\mu - k \cdot \sigma, \mu + k \cdot \sigma]$ have so little probability that they are considered to be impossible. Depending on how low we want this probability to be, we can take $k = 3$, or $k = 6$, or some other value k. So, it is reasonable to assume that the lower endpoint of this interval corresponds to t_2, i.e., that

$$\mu - k \cdot \sigma = t_2.$$

Hence, for given t_1 and t_2, once we know μ, we can determine σ.

Thus, to find the corresponding distribution, it is sufficient to find the corresponding value μ.

As this mean value μ, it is reasonable to take the result of the deterministic prediction, i.e., $\mu = t_2 + (t_2 - t_1)$. In this case, from the above formula relating μ and σ, we conclude that $\sigma = (t_2 - t_1)/k$.

Finally, an explanation. Now we are ready to explain the above empirical fact.

In the case of a single big dose, when the difference $t_2 - t_1$ is small, most of the probability – close to 1 – is located is the small vicinity of t_1, namely in the $k\sigma$ interval which now takes the form $[t_2, t_2 + 2(t_2 - t_1)]$. Thus, in this case, we have:

- (almost highest possible) probability $p \approx 1$ that the next occurrence will happen in the short-term time interval and
- close to 0 probability that it will happen in the long-term time interval.

Not surprisingly, in this case, we get:

- a better short-term effect than in the two-dose case, but
- we get much worse long-term effect.

In contrast, in the case of delayed repetition, when the difference $t_2 - t_1$ is large, the interval $[t_2, t + 2(t_2 - t_1)]$ of possible values t_3 spreads over long-term times as well. Thus, here:

- the probability p to be in the short-time interval is smaller than the value ≈ 1 corresponding to the one-dose case, but
- the probability to be in the long-term interval is larger than the value ≈ 0 corresponding to one-dose case.

As a result, for the two-dose vaccination:

- we may get worse short-term effect, but
- we get much better long-term effect,

exactly as empirically observed.

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), and HRD-1834620 and HRD-2034030 (CAHSI Includes). It was also supported by the program of the development of the Scientific-Educational Mathematical Center of Volga Federal District No. 075-02-2020-1478.

References

[1] L. Bokati, J. Urenda, O. Kosheleva, and V. Kreinovich, "Why immediate repetition is good for short-term learning results but bad for long-term learning: explanation based on decision theory", In: M. Ceberio and V. Kreinovich (eds.), *How Uncertainty-Related Ideas Can Provide Theoretical Explanation for Empirical Dependencies*, Springer, Cham, Switzerland, to appear.

[2] D. J. Sheskin, *Handbook of Parametric and Non-Parametric Statistical Procedures*, Chapman & Hall/CRC, London, UK, 2011.

Received: February 12, 2021