# Stages of human development

Luis Avitia[1], Daniel Duru[1], Eduardo Figueroa[1], Diego Reyes[1],
Leobardo Valera[2], and Vladik Kreinovich[1]
[1]Department of Computer Science, University of Texas at El Paso
500 W. University, El Paso, Texas 79968, USA
ldavitia@miners.utep.edu, dduru1@miners.utep.edu,
efigueroa5@miners.utep.edu, dareyes11@miners.utep.edu, vladik@utep.edu
[2]Computer Science and Information Technology Systems Department
El Paso Community College, 919 Hunter, El Paso, TX 79915, USA
leobardovalera@gmail.com

**Formulation of the problem.** Researchers classify post-birth human development into stages: (1) infancy and toddlershood (birth to approximately 3 years), (2) early childhood (preschool years, approximately 3-6 years), (3) middle childhood (school years, approximately 6-11 years), (4) adolescence (approximately 11-19 years), (5) early adulthood (approximately 20-40 years), (6) middle adulthood (approximately 40-65 years), and (7) late adulthood (65 years and older).

Interestingly, the last moment of each stage is approximately twice the previous one – with the exception of the last two stages. This phenomenon can be extended to all stages if we consider the last two stages as a single one. How can we explain this?

**What we plan to explain and how.** Each stage is prepared by the previous stage. In particular, this means that the characteristics of the next stage – e.g., its last moment – are uniquely determined by the previous stage. We can describe this dependence as $t_{n+1} = f(t_n)$, where $t_i$ is the last moment of the $i$-th stage and $f(t)$ is some function that characterizes this dependence. In particular, the observed moments correspond to $f(t) = 2t$.

What we will explain is that $f(x)$ should be a linear function $f(x) = a \cdot x$ for some $a$. We are still not able to explain why $a = 2$.

**The main idea behind our explanation.** In principle, we can measure time by using different units – seconds, minutes, hours, days, etc. For many real-life processes, there is no fixed measuring unit, the formulas should not change if we use a different unit. When we change a measuring unit to a one that is $c$ times smaller, then all the numerical values of moments $t_i$ gets multiplied by $c$. For example, 2 minutes $=$ $(60 \cdot 2 =)$ 120 seconds. Thus, the above requirement takes the following form: if in the original units, we had $s = f(t)$, then in the new units, we should have $s' = f(t')$, where $s' = c \cdot s$ and $t' = c \cdot t$. Let us show that this natural requirement indeed leads to $f(t) = a \cdot t$.

**Resulting explanation.** Substituting the values $s' = c \cdot s$ and $t' = c \cdot t$ into the equality $s' = f(t')$, we conclude that $c \cdot s = f(c \cdot t)$. Here, $s = f(t)$, so $c \cdot f(t) = f(c \cdot t)$ for all $c > 0$ and $t > 0$. In particular, for any $x > 0$, we can take $c = x$ and $t = 1$ and get $f(x) = a \cdot x$, where $a \stackrel{\text{def}}{=} f(1)$. This explains the above empirical formula for $f(x)$.

*Comment.* Similar scale-invariance explains the use of so-called ReLU activation function $s_0(z) = \max(0, z)$ in deep learning. To be precise, similar scale-invariance requirements lead to the piece-wise linear function $s(z) = a_+ \cdot z$ for $z \geq 0$ and $s(z) = a_- \cdot z$ for $z < 0$, for some constants $a_-$ and $a_+$. This function in linearly related to ReLU, e.g., $s(z) = a_- \cdot z + (a_+ - a_-) \cdot s_0(z)$, so in a neural network that combines the use of activation function and linear transformations, its use is equivalent to using ReLU.