Sigmoid functions in reliability based management

Tamás Jónás

Abstract

This paper introduces the theoretical background of a few possible applications of the so-called sigmoid-type functions in manufacturing and service management.

An extended concept of reliability, derived from fuzzy theory, is discussed here to illustrate how reliability based management decisions can be made consistent, when handling of weakly defined concepts is needed.

I demonstrate how performance growth can be modelled using an aggregate approach with support of sigmoid-type functions.

Suitably chosen parameters of sigmoid-type functions allow these to be used as failure probability distribution and survival functions. If operation time of an item has a given sigmoid-type failure probability distribution function, then its hazard function is proportional to the failure probability distribution function. Furthermore, this hazard function can be a model of the third part of the bathtub failure rate curve.

Keywords

Performance growth · sigmoid function · aggregate performance · performance growth · extended concept of reliability · survival function · hazard function

Tamás Jónás

Flextronics International Ltd.,, 1183 Budapest , Hangár u. 5-37, Hungary
e-mail: tamas.jonas@hu.flextronics.com

1 Introduction

A sigmoid function is a mathematical function that produces a curve having an "S" shape, and is defined by the

$$\sigma_{\lambda,x_0}(x) = \frac{1}{1 + e^{-\lambda(x-x_0)}} \qquad (1)$$

formula. The sigmoid function is also called sigmoidal curve [1] or logistic function. The interpretation of sigmoid-type functions – I use here – is that any function that can be transformed into $\sigma_{\lambda,x_0}(x)$ through substitutions and linear transformations is considered as a sigmoid-type one. There are several well-known applications of sigmoid-type functions. A few examples are: threshold function in neural networks [2], approximation of Gaussian probability distribution, logistic regression [3], or membership function in fuzzy theory [4].

My objective is to conclude hypotheses on how this function family is applicable in certain areas of reliability based manufacturing and service management, along with brief interpretations and demonstrations of these possible applications. Besides the (1) generic form, different other forms of the sigmoid function such as $P(x)$, $\mu(m_R)$, $F_{\lambda,t_0}(t)$ and $R_{\lambda,t_0}(t)$ with different parameters will be used here. These different forms are different appearances of the same function, and the notations always fit to the notations that are commonly used in the fields of particular applications.

2 Modelling performance growth

The manufacturing as well as the service processes can be characterized by various indicators and metrics, which are usually functions of several process variables, parameters and constants. The overall performance of a process depends on its inputs, and it is common that finally, there is one aggregated indicator or metric used to characterize the overall performance. These kind of aggregate indicators are commonly associated with some financial metrics, and so whenever a new process is being introduced its financial performance can be monitored through the chosen aggregate performance indicator. Certainly, the ultimate goal is to find the highest performance resulting input set as quickly as possible. However, in reality, the manufacturing and service processes are too complex, with a large num-

ber of input factors, and commonly, with unknown dependencies and interactions among them. That is, searching for dependencies between the input set and the aggregate performance is cumbersome and works with difficulties in most of the practical cases. Instead of trying to handle many inputs and outputs, the approach shown here is using one independent variable that is proportional to and so represents all those efforts that contribute to the aggregate performance growth. For example, we may consider the time spent on process development and improvement as an aggregate input variable, and so we can look at the aggregate performance growth as function of the so-interpreted time variable.

2.1 The nature of performance growth

Let x be such an independent input variable that is proportional with the performance development and improvement efforts. Furthermore, let $P(x)$ note the aggregate performance as function of x, P_i the initial value and P_t the target value of $P(x)$. With other words, we look at the growth of aggregate performance as function of x, providing that $P(x)$ increases from P_i to P_t, and P_t represents the level of operational performance corresponding to the targeted financial results.

Numerous practical observations confirm that the same Δx effort increase results different $\Delta P(x)$ performance increases, depending on the actual level of performance, at which the efforts were made. It may be assumed that when $\Delta P(x)$ is at a low level and is close to P_i the speed of growth is low, namely, the early efforts do not yield much of improvement. As the performance increases, the speed of its growth increases as well. It may be explained so that as the improvement and development efforts result higher and higher level of process specific knowledge and skills, the impact of the same sized every new effort results greater performance increase. This tendency, however, is only valid until a certain level of performance. Although the performance increases as the efforts increase, but after a certain time, the growth slows down as the new efforts result low-rate increase in process specific knowledge and skills. Finally, the performance gets close to an upper limit. If the P_t target value of $\Delta P(x)$ is set to this upper limit, then it can be said that growth speed of $\Delta P(x)$ decreases when $P(x)$ is near P_t.

Fig. 1 illustrates the growth speed of performance in its different ranges, and shows how a small Δx effort increase results low-rate performance increase when $P(x)$ is near P_i or P_t, as well as shows greater slope of $P(x)$ when it is more distant both from P_i and P_t. This illustration assumes that the chosen Δx is small enough to assume linear relationship between x and $P(x)$ in the Δx intervals.

Thus, my assumption is that the nature of performance growth is so that its speed is proportional to the $P(x) - P_i$ and $P_t - P(x)$ differences. Formally,

$$\Delta P(x) = \lambda^*\big[P(x) - P_i\big]\big[P_t - P(x)\big]\Delta x$$

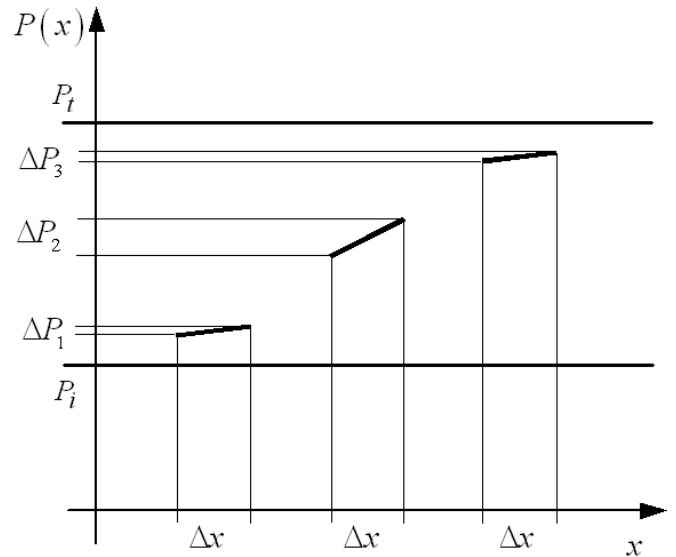$$\frac{\Delta P(x)}{\Delta x} = \lambda^*\big[P(x) - P_i\big]\big[P_t - P(x)\big] \tag{2}$$

Fig. 1. Growth speed of performance in its different ranges

where $\lambda^* > 0$ is a process specific proportionality coefficient. Turning into infinitesimal quantities results the following differential equation.

$$\frac{dP(x)}{dx} = \lambda^*\big[P(x) - P_i\big]\big[P_t - P(x)\big] \tag{3}$$

Eq. 3 is known as logistic equation and is also used as a model of population growth [5]. Population models using the logistic growth can be found in Murray's book [9] and the book by Clark [10] introduces its applications in economics. Solving this equation results the

$$P(x) = \frac{P_t e^{\lambda(x-a)} + P_i}{1 + e^{\lambda(x-a)}} \tag{4}$$

function, where $\lambda = -\lambda^*(P_i - P_t) = \lambda^*(P_t - P_i)$ is a positive number. If $a = x_0$, $P_i = 0$, and $P_t = 1$, then

$$P(x) = \frac{e^{\lambda(x-x_0)}}{1 + e^{\lambda(x-x_0)}} = \frac{1}{1 + e^{-\lambda(x-x_0)}}. \tag{5}$$

It means that function (4) is a sigmoid-type function.

2.2 Attributes of the performance growth function

The $P(x)$ function derived above has four parameters: λ, a, P_i, and P_t. Interpretation of these parameters and the basic analytical properties of $P(x)$ are introduced in this subsection.

Derivative
Derivative of $P(x)$ is

$$\frac{dP(x)}{dx} = \frac{\lambda e^{\lambda(x-a)}(P_t - P_i)}{(1 + e^{\lambda(x-a)})^2} \tag{6}$$

Monotonicity and limits
As $P_t - P_i$ and λ are positive, the derivative is positive as well, and so $P(x)$ is an increasing function.

$$\lim_{x \to -\infty} P(x) = P_i \tag{7}$$

and

$$\lim_{x \to \infty} P(x) = P_t \qquad (8)$$

that is $P(x)$ is an increasing function from P_i to P_t.

Symmetry and inflection point

It can been seen that $P(x)$ has its only one inflection point in the $(a, \frac{P_i+P_t}{2})$ point, wherein it changes its shape from convex to concave. This point is also the symmetry center of the $P(x)$ curve.

Role of parameter λ

The (6) derivative in a is

$$P'(a) = \frac{\lambda}{4}(P_t - P_i). \qquad (9)$$

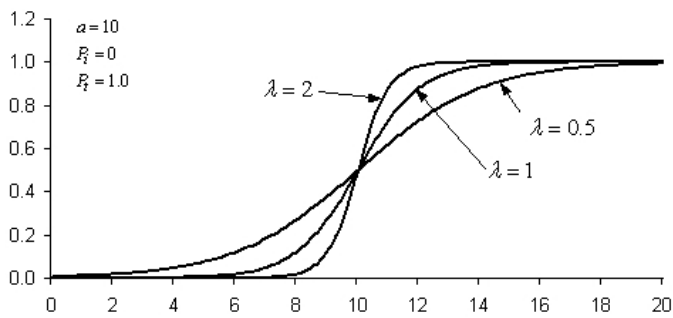It means that role of λ relates to the speed of change, since slope of the curve in a is proportional to λ.

Fig. 2. Role of λ

Impact of parameter a

As it is shown in Fig. 2, the $P(x)$ curve has its only one inflection point in a. Graphically, it means that parameter a determines the point where the "S" curve takes its place along the abscissa axis. The $(a, \frac{P_i+P_t}{2})$ point is the one, around which performance growth is the fastest. Fig. . 4 shows the effect of parameter a, when all the other parameters are kept unchanged.
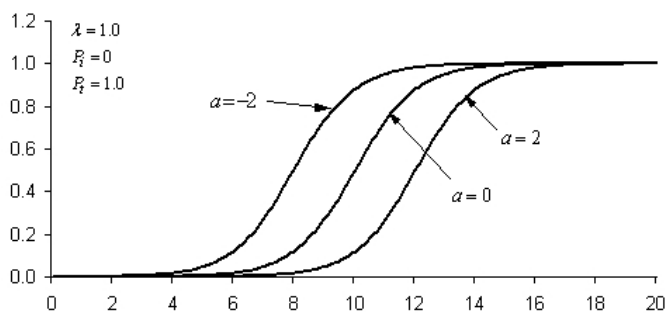

Fig. 3. Parameter a determines the place of curve along the x-axis

The function curve

Considering that P_i and P_t stand for the initial, and target performance levels respectively, the function curve with indication of meaning of its parameters is in Fig. 4.
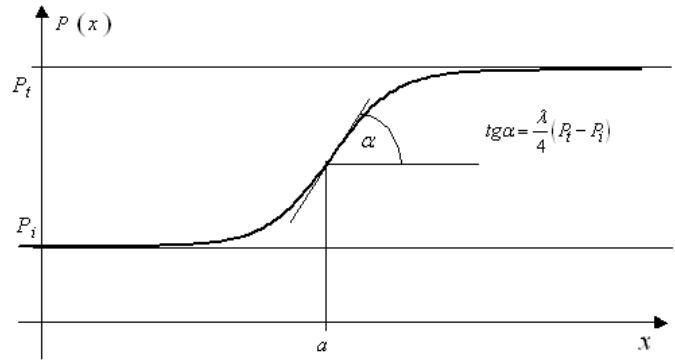

Fig. 4. A generic performance growth function with its four parameters

2.3 Sigmoid-type functions as possible models of performance growth

Hypothesis 1

Let x represent the time effectively spent on the introduction of a new manufacturing or service process, and let $P(x)$ be an aggregated performance metric used to characterize the goodness of this process. My assertion is that $P(x)$ is a sigmoid-type function shown in (4), with the λ, a, P_i, and P_t parameters, moreover each process introduction has its unique set of these parameters.

Providing that this assertion is valid, the model would be applicable predicting the performance growth, if λ, a, P_i, and P_t are determined once. On the grounds of these, I hypothesize the following:

Hypothesis 2

The sigmoid-type functions that have the form (4) and the λ, a, P_i, and P_t parameters, can be used as control tools in new product or service introductions so that the time dependent growth of an aggregated process performance characteristic is measured against the corresponding time dependent values of a suitably parameterized sigmoid-type function. Suitable setting of the parameters means that the uniquely determined, best fitting resulting parameters are chosen as described in Hypothesis 1.

Validity of these assertions and a much deeper investigation on what factors, conditions and circumstances drive the parameters are subjects of further research. Assuming validity of these two hypotheses allows us to proceed like in the next example. Imagine that a new service process has been introduced, and the time dependent values of a chosen aggregated performance metric have been collected during the introduction. It means that a set of (performance value, time) type ordered pairs is available, and based on these, the λ, a, P_i, and P_t values of the best fitting sigmoid-type function are determined. Thus, we have a model that describes the performance growth observable during the introduction of this particular process. Later on, when the same process has to be introduced under the same condi-

tions and circumstances again, then it may be assumed that its performance growth will follow the same behavior curve as performance growth of the firstly introduced process did. Thus, the determined performance growth function can be used for target setting, and the real growth of performance can be compared to the predicted values. It may be worrying and may sound impractical that existence of "same conditions and circumstances" was assumed in this example. The concept of sameness used here does not mean a literal identity since use of that would be really far from the practice. Application of a fuzzy concept of sameness can resolve this problem, and I would like to study this area more in details. Fuzzy concepts are briefly discussed in the next section.

3 Extension of concept of reliability

When the company processes are interpreted so that besides the manufacturing and service processes, the support functions needed to a proper internal operation are taken into consideration as well, then characterizing reliability of a process may require handling of weakly defined concepts. For example, when may we consider a human resource selection process reliable? In this case, the "reliable process" notion is a weakly defined one, and that is why it is difficult to handle. Even if some measurable metrics can be assigned to this process, drawing a sharp borderline between the reliable and unreliable domains may be unsuitable. In such cases the simple, metrics based categorization into different domains may jeopardize the real consistency of decision-making. On the other hand, treating the "reliable" and "unreliable" domains as fuzzy concepts, and having a suitable membership function that, in fuzzy manner, can decide which domain a particular process belongs to, would result a more practical decision-making. Bellman and Zadeh in [5] say: "By decision-making in a fuzzy environment is meant a decision process in which the goals and/or the constraints, but not necessarily the system under control, are fuzzy in nature. This means that the goals and/or the constraints constitute classes of alternatives whose boundaries are not sharply defined."

For example, let m_R be a reliability metric of a process, and our task is to decide if this process is reliable enough or not.

3.1 The traditional approach

Following the traditional way, we would define a sharp m_T limit for m_R and base the decision on comparing the particular m_R^* value of m_R to the m_T limit. Formally, we would use the following $D(m_R)$ decision function:

$$D(m_R) = \begin{cases} 0 & \text{if } m_R < m_T \\ 1 & \text{if } m_R \geq m_T \end{cases} \tag{10}$$

where the 0 and 1 logical values correspond to the unreliability and reliability respectively.

3.2 A possible fuzzy approach

If there are uncertainties influencing the values of m_R, and that is why the reliable and unreliable domains are not sharply disjoint, then the examined process may be considered as a reliable one even if m_R is less than m_T, but is close to m_T, and on the contrary, the process may be considered as an unreliable one even if m_R is greater than m_T, but is close to m_T. The decision making can be supported by the

$$\mu(m_R) = \frac{1}{1 + e^{-\lambda(m_R - m_0)}} \tag{11}$$

membership function, where λ and m_0 have the same roles as λ and a have in sub-section 2.2. It is important to see that these two parameters are process specific ones, and in this manner, are unique properties of the process examined. In this case, the decision-making process works so that for the particular m_R^* reliability level the $\mu(m_R^*)$ truth value is calculated. This truth value reflects how much valid the "process is reliable" statement is. Afterwards, the so-calculated $\mu(m_R^*)$ truth value is compared to the pre-defined T_L truth limit that represents the threshold, which the decision-making is based on. We may accept that the process is reliable, if

$$\mu(m_R^*) \geq T_L = \mu(m_L) \tag{12}$$

The difference between the traditional and fuzzy approaches is that while the traditional decision is based on two values of the traditional logic, the fuzzy approach compares two truth values.
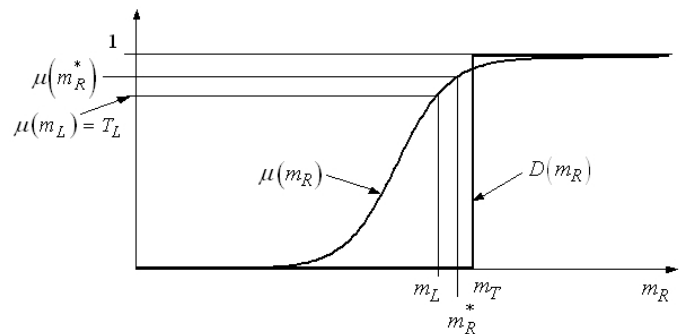

Fig. 5. Reliability based decision making in traditional and fuzzy manner

The example in Fig. 5 demonstrates how these two approaches work. In this example, the particular m_R^* value of the chosen reliability metric is less than the m_T limit, and so the process in unreliable in traditional manner. On the other hand, as the difference between m_R^* and m_T is small, it may be assumed that this difference is just caused by the uncertainty of the m_R metric, and therefore, it is an option to accept that the process is reliable. If the T_L truth limit is set as in Fig. 5, then $\mu(m_R^*) \geq T_L = \mu(m_L)$, and the reliability of the process can be accepted in fuzzy manner.

Certainly, the output of this fuzzy-logic based decision making depends on the $\mu(m_R)$ function, and the T_L limit, and so it is key how these are defined and set up in a concrete, practical

case. Further investigations and development of models for particular applications are subjects of my future research activities. Here I give just a brief explanation on how the λ and m_0 parameters of $\mu(m_R)$, and the T_L limit impact the decision. I study the impact of parameters separately, so that I examine the effect of changing only one factor at the same time.

Changing m_0 results the $\mu(m_R)$ curve shifting along the m_R axis. If m_0 decreases, while λ and T_L are fixed, the curve is getting shifted to left, and $\mu(m_R)$ gives greater values for the same m_R inputs. It means that a lower value of the m_R reliability metric gets higher truth value than its original truth value was, and so the "trust" in process reliability strengthens. With other words, we would accept that the process is reliable, even if its reliability metric has decreased. Similarly, as m_o increases, the decision making mechanism becomes more strict, and values of m_R that - based on their truth value - originally belonged to the reliable domain, fall into the unreliable domain.

As λ determines the slope of the $\mu(m_R)$ curve in the $(m_0, \frac{1}{2})$ point, changing λ impacts the sharpness of border between the unreliable and reliable domains. If $m_0 = m_R$ and λ tends to infinity, then $\mu(m_R) = D(m_R)$.

T_L represents our expectation in terms of truth value that we require the process to meet in order to consider it being reliable. T_L can also be considered as the indulgence level of decision maker.

Hypothesis 3

Reliability based decision-making situations, which are uncertain due to the lack of sharply defined reliability domains, can be handled through sigmoid-type functions. In such a case, a sigmoid-type function with expediently chosen parameters can be used as a truth function, whose independent argument is a reliability metric of the examined process, and output is a truth value between 0 and 1. This truth value measures the validity of the statement that the examined process is reliable.

4 Sigmoid-type functions as survival functions
The

$$\sigma_{\lambda,x_0}(x) = \frac{1}{1 + e^{-\lambda(x-x_0)}} \tag{13}$$

sigmoid-type function meets the criteria of a probability distribution function as

- $\lim\limits_{x \to -\infty} \sigma_{\lambda,x_0}(x) = 0$

- $\lim\limits_{x \to \infty} \sigma_{\lambda,x_0}(x) = 1$

- $\sigma_{\lambda,x_0}(x)$ is monotonously increasing

- $\sigma_{\lambda,x_0}(x)$ is continuous from left side.

Function (13) is known as logistics distribution function. Based on this

$$F_{\lambda,t_0}(t) = \frac{1}{1 + e^{-\lambda(t-t_0)}} \tag{14}$$

is a possible failure probability function[1], and the

$$R_{\lambda,t_0}(t) = 1 - F_{\lambda,t_0}(t) = 1 - \frac{1}{1 + e^{-\lambda(t-t_0)}} = \frac{e^{-\lambda(t-t_0)}}{1 + e^{-\lambda(t-t_0)}} \tag{15}$$

is the corresponding survival function. If the τ lifetime of an element[2] has the $F_{\lambda,t_0}(t)$ probability distribution, then the probability that $\tau \geq t$ is $R_{\lambda,t_0}(t)$:

$$P(\tau \geq t) = R_{\lambda,t_0}(t) = \frac{e^{-\lambda(t-t_0)}}{1 + e^{-\lambda(t-t_0)}} \tag{16}$$

The $R_{\lambda,t_0}(t)$ function is a sigmoid-type function as multiplying the numerator and denominator of (16) by $e^{\lambda(t-t_0)}$ results

$$R_{\lambda,t_0}(t) = \frac{e^{-\lambda(t-t_0)}}{1 + e^{-\lambda(t-t_0)}} = \frac{1}{1 + e^{\lambda(t-t_0)}} = F_{-\lambda,t_0}(t) \tag{17}$$

that is equal to (14), if λ is negative. It means that changing the sign of λ in the $F_{\lambda,t_0}(t)$ failure probability distribution function yields the corresponding $R_{\lambda,t_0}(t)$ survival function.

4.1 Properties of $R_{\lambda,t_0}(t)$
Considering properties of the $F_{\lambda,t_0}(t)$ failure probability function, the $R_{\lambda,t_0}(t)$ survival function has the following attributes:

- $R_{\lambda,t_0}(t)$ is monotonously decreasing from 1 to 0

- $R_{\lambda,t_0}(t)$ is changing its shape from concave to convex in t_0

- $R_{\lambda,t_0}(t)$ has angular coefficient of $-\dfrac{\lambda}{4}$ in the $(t_0, \dfrac{1}{2})$ point.

The t_0 parameter represents a kind of a threshold for the τ lifetime as the $P(\tau \geq t)$ likelihood is changing from 1 to 0 in neighbourhood of t_0, and the speed of this change is determined by λ.

If $t \ll t_0$ then $P(\tau \geq t) \approx 1$, that is the probability that the τ lifetime is greater than t is approximately 1. Similarly, if $t \gg t_0$ then $P(\tau \geq t) \approx 0$, or with other words, the element will very likely fail, if it operates considerably longer than t_0. In neighbourhood of t_0, the λ parameter drives the speed of transition from 1 to 0. If λ is infinitely large, the transition is infinitely short, that is in this case the $R_{\lambda,t_0}(t)$ survival function belongs to an element that operates without any failure till t_0, and immediately fails as its operation time reaches t_0.

4.1.1 Expected lifetime
One important attribute of the lifetime of an element is its expected value. If the τ lifetime of an element has the $F(t) = P(\tau < t)$ failure probability distribution and the $F'(t) = f(t)$ density function, then expected lifetime[3] of the element is

$$E(\tau) = \int\limits_0^\infty t f(t) \mathrm{d}t \tag{18}$$

---

[1] Certainly, when this function is used as a failure probability distribution function, then its domain of variability is positive.

[2] From this point onwards, wherever "element" is referred to, the same statement is valid for a system as well.

[3] Expected lifetime is also called the mean time to failure, expected time to failure, or average life.

Derivation of the expected value and other properties of logistics distribution can be found in the book by Johnson, Kotz and Balakrishnan [11]. The expected value with my notation is

$$E(\tau) = \frac{1}{\lambda} \ln(1 + e^{\lambda t_0}). \qquad (19)$$

4.1.2 The hazard function

Generally,

$$\frac{R(t) - R(t + \Delta t)}{R(t)} \qquad (20)$$

is the conditional probability that an item with $R(t)$ survival function fails in the $[t, t + \Delta t]$ time interval given that it has not failed till time t. If Δt is infinitely small, then (20) equals to $\lambda(t)\Delta t$, where

$$\lambda(t) = \frac{-R'(t)}{R(t)} = \frac{F'(t)}{R(t)} = \frac{F'(t)}{1 - F(t)} \qquad (21)$$

[7] is the so-called hazard function[4].

If $R(x) = R_{\lambda,t_0}(t)$, then

$$\frac{-R'(t)}{R(t)} = \frac{-R'_{\lambda,t_0}(t)}{R_{\lambda,t_0}(t)} = \frac{F'_{\lambda,t_0}(t)}{R_{\lambda,t_0}(t)} = \frac{F'_{\lambda,t_0}(t)}{1 - F_{\lambda,t_0}(t)} \qquad (22)$$

Derivation of $F_{\lambda,t_0}(t)$ results

$$\left[1 - \frac{1}{1 + e^{-\lambda(t-t_0)}}\right] \frac{1}{1 + e^{-\lambda(t-t_0)}} \qquad (23)$$

that is

$$F'_{\lambda,t_0}(t) = \lambda\left[1 - F_{\lambda,t_0}(t)\right] F_{\lambda,t_0}(t) \qquad (24)$$

and so the hazard function is

$$\lambda(t) = \frac{F'_{\lambda,t_0}(t)}{1 - F_{\lambda,t_0}(t)} = \lambda \frac{\left[1 - F_{\lambda,t_0}(t)\right] F_{\lambda,t_0}(t)}{1 - F_{\lambda,t_0}(t)} = \lambda F_{\lambda,t_0}(t) \qquad (25)$$

Looking at $\lambda(t)$, it can be seen that it is proportional to $F_{\lambda,t_0}(t)$, and since $F_{\lambda,t_0}(t)$ converges to 1 as t tends to infinity, $\lambda(t)$ converges to λ. This conclusion enables using this $\lambda(t)$ in the third part of the bathtub failure rate curve [8]. My assumption is that the sigmoid-type function introduced in (15) can be used as an approximating function of the survival function of certain components. Statistical verification of this conjecture is a subject of my further research plans.

5 Summary

The following three possible applications of sigmoid-type functions in reliability based management were introduced here.

• There are well known applications of the logistic equation for growth modelling in different areas such as biology, chemistry or economics. In this paper, application of the logistic equation and sigmoid functions $\left(P(x)\right)$ as growth models for aggregate performance of manufacturing and service processes were introduced.

• When reliability or unreliability of a process cannot be unambiguously judged, sigmoid functions as membership functions $\left(\mu(m_R)\right)$ can support drawing appropriate conclusions on the process reliability. The introduced approach represents an extended concept of reliability.

• Sigmoid functions are possible failure probability distribution $\left(F_{\lambda,t_0}(t)\right)$ and survival functions $\left(R_{\lambda,t_0}(t)\right)$. In this article, I have shown that the hazard function is proportional to the failure probability distribution function if the failure probability distribution function is a sigmoid one. This result suggests possible applications of sigmoid functions for modelling the third part of the bathtub failure rate curve.

References

1 **von Seggern D**, *CRC Standard Curves and Surfaces with Mathematics*, FL: CRC Press, 2007. 2nd ed. Boca Raton.

2 **Mitchell TM**, *Machine Learning*, WCB-McGraw-Hill, 1997.

3 **Hosmer D, Lemeshov S**, *Applied Logistic regression*, Wiley & Sons, New York, 1989.

4 **Dombi J**, *Membership function as an evaluation*, Fuzzy Sets and Systems **35** (1990), 1-22, DOI 10.1016/0165-0114(90)90014-W.

5 **György Cs**, A biológiai szabályozás, Medicina Könyvkiadó, Budapest, 1978, pp. 360-362.

6 **Bellman RE, Zadeh LA**, *Decision Making in a Fuzzy Environment*, Washington D. C., May 1970. NASA CR-1594.

7 **Balogh A, Dukáti F, Sallay L**, *Minőségellenőrzés és Megbízhatóság*, Műszaki Könyvkiadó, Budapest, 1980.

8 **Hnatek ER**, *Practical Reliability of Electronic Equipment and Products*, CRC Press, 2003.

9 **J. D. Murray**, *Mathematical Biology*, Springer-Verlag, New York, 1989.

10 **Clark C**, *Mathematical Bioeconomics: The Optimal Management of Renewable Resources*, Wiley Interscience, New York, 1990. second edition.

11 **Johnson NL, Kotz S, Balakrishnan N**, *Continuous Univariate Distributions*, Wiley Series in Probability and Statistics **2** (1995), 113-153.

---

[4]The hazard function is often called conditional failure rate function, failure rate function, or hazard rate.