PMWorldToday

Published in *PM World Today* – March 2010 (Vol XII, Issue III)

PM WORLD TODAY – FEATURED PAPER – MARCH 2010

Nonlinear and Random Accomplishment Time of Human Actions

By Pavel Barseghyan, PhD

Abstract

Human activities can be represented as a set of parallel and sequential human actions. In other words, separate actions of humans are the elements of their activities.

Each human action can be characterized by its difficulty and duration. Analyzing the usual sequence of everyday human actions can be seen that they have a random nature and can be described by probabilistic methods.

Besides, quantitative analysis of individual human actions indicates that the relationships between action duration and action difficulty parameters have nonlinear nature.

Another important property of human actions is that part of their parameters can be measured relatively easy. For this reason, quantitative study of individual human actions is crucial for the establishment of the basic mathematical theory of human labor.

Having the quantitative picture of the individual human actions, through their parallel and serial combinations one can create a general quantitative picture of the whole human labor.

Such an approach provides a rich opportunity for applying the methodology of systems analysis in the field of human labor. Under this approach, human activities are considered as a system whose elements are the individual actions of people.

In addition, any delay in the work of people and the associated risks are closely linked with the quantitative characteristics of individual human actions. In this sense, it is important to derive in an analytical way and justify the distribution functions of the duration of human activities on the basis of the fundamental mathematical description of the individual human actions.

This approach to determining the schedule risk of people's work shows that the choice of a particular type of risk function cannot be arbitrary, as it takes place currently, because it is closely linked with the essence of the specific problems under study.

This work is devoted to the study of quantitative models of human actions in order to reflect their nonlinear and stochastic nature.

The work consists of several parts. In the first part with the help of a statistical experiment are derived probabilistic characteristics of the duration of human actions. The differential equation to describe the dynamics of the completion of human actions is derived. The notion of intensity of the completion of individual human actions is introduced. Different distribution functions of human action duration are discussed. Depending on the circumstances, these distribution functions can have either finite or infinite variance, thus explaining the emergence and mechanisms of fat tail distributions.

Key words: Duration of human actions, nonlinear and stochastic behavior, intensity of the completion of human actions, probability of the completion of human actions, fat tail distributions, Weibull distribution, Pareto distribution.

Introduction

The relative easiness of measurements of the duration of human actions is very important for the development of the mathematical theory of human activity. Based on the human action duration statistics it is possible to build mathematical models of single actions. Combining that with the fact that human activity of any size, including projects, can be represented in the form of mixed networks of the parallel and serial human actions, it is possible to build mathematical models of the whole human activity as a totality of human actions.

The development of the quantitative project management encounters with a number of fundamental concerns and the problem of measurements is only one of them. Along with the unresolved problems of project data mining and project change management in the course of project execution, the measurement problems are the main obstacles that impede the establishment of scientifically sound mathematical theory of project management.

To analyze this problem it suffices to consider the uncertainties associated with measurements of people's productivity. It is known that human productivity is closely linked to the difficulty of the problems being solved and the number of people working together [1]. In these conditions it is not clear what it does mean to measure the people's productivity.

Before any measurement it is necessary to clarify what should be measured, performance of a single person, group of people or of an entire project team? It is also unclear when to make measurements, at the beginning of the project, when people have low productivity, or in the final stages of the project, when people are working overtime trying to catch up [2].

From this perspective it is important to know what kind of measurements we need to perform in the course of human work, when we need to do that and at what organizational level these measurements have to be done?

It is especially important to analyze the problem of measurement in hierarchical systems, because it has direct impact on the system development methodology.

In this sense it is important to use the accumulated experience from other fields of science and engineering where the roles of experiments and mathematical theories have been established long time ago.

The typical example is the field of reliability engineering where there is a clear division between experimentation at the element level for defining failure characteristics of elements and pure probabilistic modeling of structures at system level [6].

A common rationale for this approach is that as a rule, the number of types of elements in the systems is limited, that allows creating libraries of their characteristics.
Practically, this kind of approach cannot be implemented at the systems level, because if the number of types of elements in the system is limited, the same cannot be said for the system.

The reason is that with the finite number of types of elements one can build an unlimited number of systems, which makes it impossible to develop the libraries of systems. For this reason, the contemporary system building methodology is based on the simulations of systems that are built using the specified libraries of elements. Therefore it is important to concentrate experiments at the element level of the hierarchy.

The same methodology is applicable for the quantitative study of human activities including project type of activities. This means that according to this approach quantitative studies of human activities have to be started with the quantitative studies of the separate human actions, including their measurements and subsequent processing of data. In such approach human activity acts as a system, and the separate actions of people serve as the elements of that system. The next step of this kind of study is the derivation of the characteristics of human activities considering them as a totality of separate human actions with various parallel-serial structures.

Analysis of measurement problems in the field of project management must take into account another very important circumstance. In fact carrying out pure experiments in the classical sense of the theory of experiments in the area of project management is impossible. For this reason, the project databases are created by collecting information on the projects. This means that from the standpoint of the theory of experiments, these data are the results of a passive experimentation, which in turn makes it impossible the direct usage of regression analysis for the mining of this kind of collected data.

Relationship between the duration of human actions and their difficulty

The functional relationship between the duration of human action and its difficulty is addressed in [3] and [4]. This functional relationship is presented in Fig.1, which shows the nonlinear nature of this dependency. From this picture it can be seen that if the difficulty of human action is approaching to the upper limit D_w inherent to the individual, the duration of action increases infinitely.

Human activities can be represented as a sequence of individual actions, each of which has a random difficulty, and therefore a random duration too. In the ideal case of homogeneous human actions the sequence of his or her actions can be approximated by a diffusion process around some average value D_{av}. This diffusion process is shown in Fig.1 in the form of the distribution density function $\psi(D)$. Passing through the nonlinear system of a human, this stochastic process is transformed into another stochastic process of the durations of actions, which is represented in the Fig.1 in the form of the distribution density function $f(\theta)$. Due to the nonlinear nature of this transformation the output density functions $f(\theta)$ can have an asymmetric view and a heavy tail.

This also means that the processes with different average values of the difficulties of human actions will produce different output distribution functions of the duration of actions. At the same time, under certain conditions, actions associated with the same average difficulty might be considered as a homogeneous flow of durations of human actions.

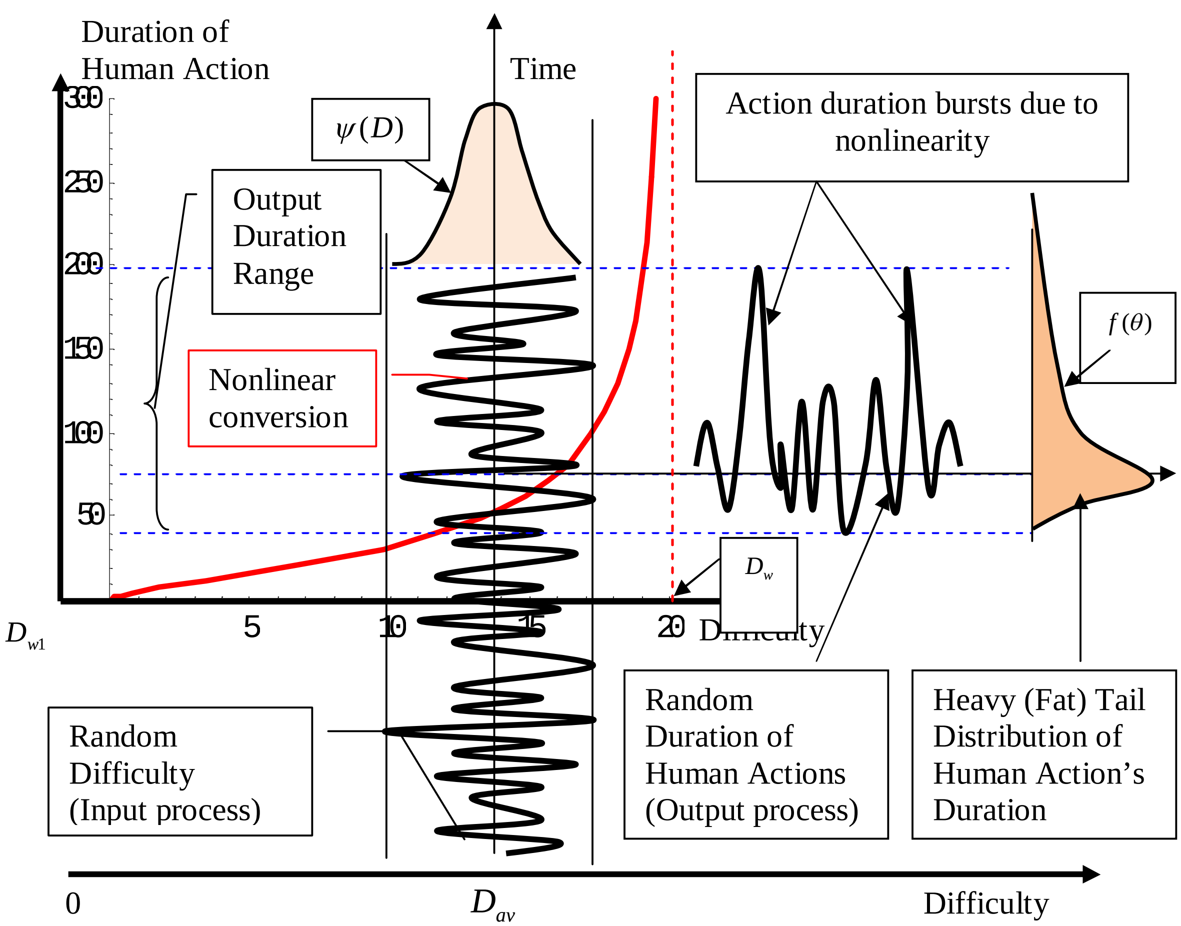


Fig.1. Nonlinear transformation of the action's difficulty into the action's timing

But in all these cases must be taken into account the fact that for different values of the average difficulty of human actions output flows of action durations may differ significantly from each other (Fig.2).

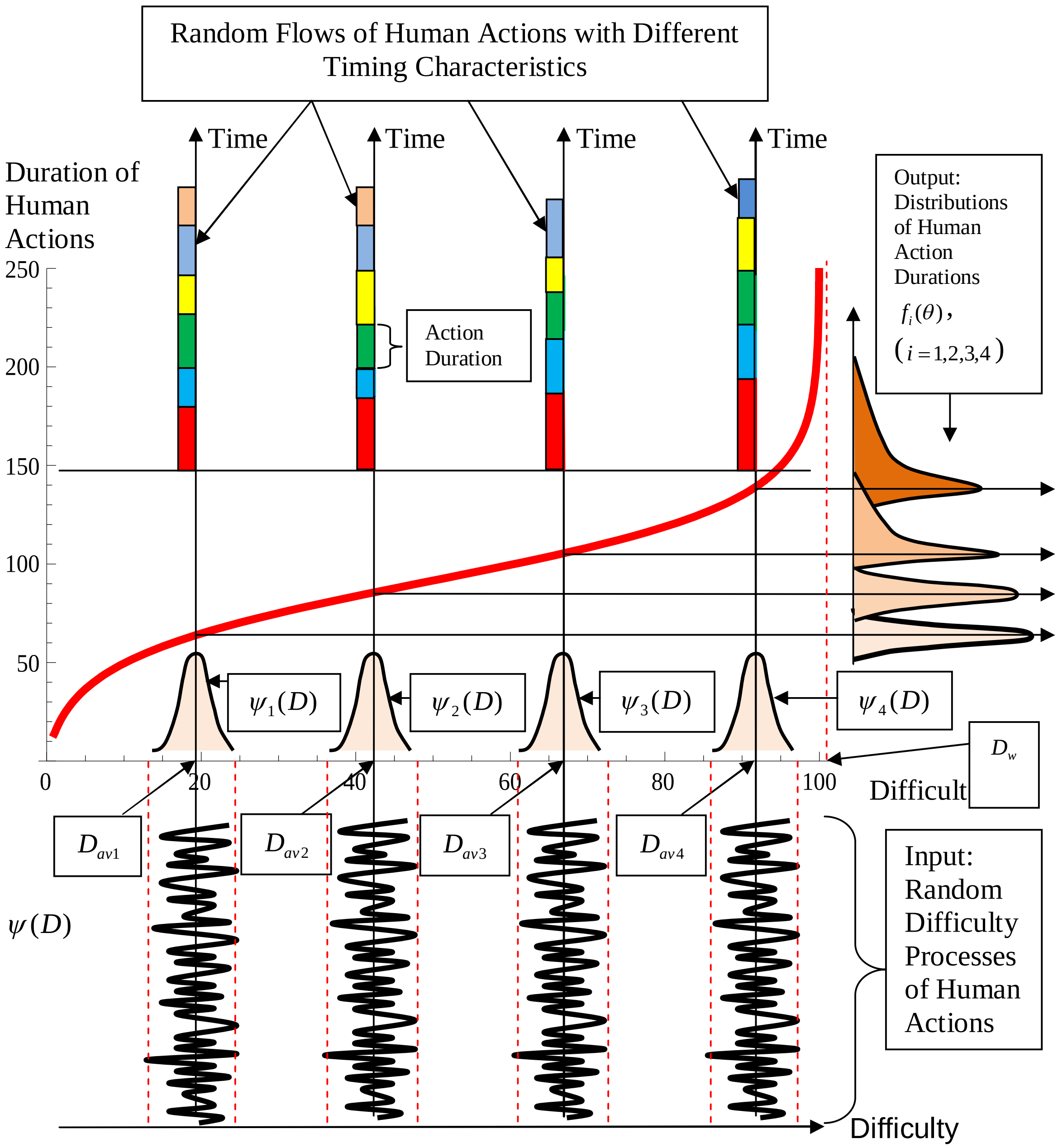


Fig.2 Transformation of four random difficulty processes into the random timing processes of human actions

This difference is due to the fact that an increase in the average difficulty of human actions increases the likelihood that the output flow of action durations will contain a larger number of durations with very large or infinite variance. But this is only one of the aspects of the problem of analysis of stochastic durations of human actions. Another aspect of this problem is associated with the stochastic nature of human productivity.

This circumstance changes situation sharply, because, as it shown in [3], under the influence of these two stochastic processes noticeable delay of human actions may occur even at the very low level of difficulty of the human actions.

Adding to this the fact that the productivity of people and the difficulty of problems to be solved are closely linked, it becomes clear that progress in this area requires a multifaceted approach to their quantitative description.

Besides the study of the functional relationship between human productivity and the difficulty of tasks must take into account, that the increase in difficulty of the problems leads to the nonlinear increase of the variance of human productivity.

Another important factor that needs to be taken into account in the mathematical modeling of such processes is the direct influence of learning processes on human productivity. This factor can be taken into account considering the upper limit of difficulties D_w as a function of time.

Analysis of the homogeneous flows of the sequential human actions

Consider the sequence of actions of one person on the time axis. As seen from the Fig.3, each action has a beginning, end and duration.

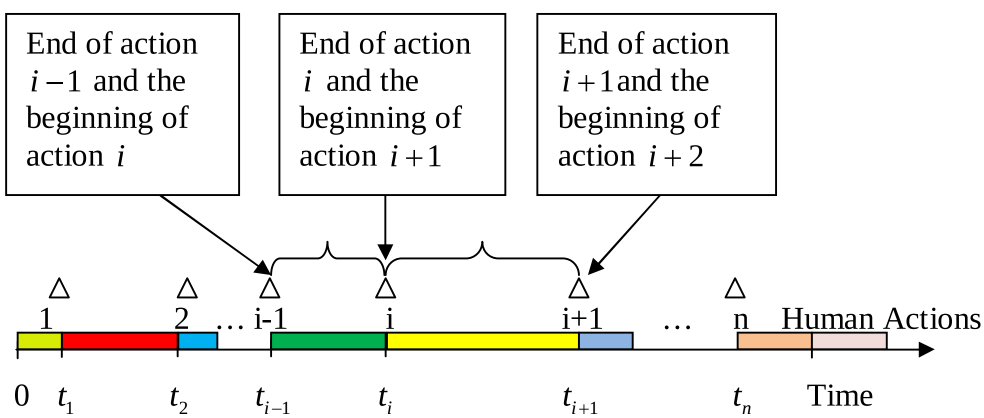


Fig.3 Sequence of actions of an individual

In order to analyze the successive actions of a human, let's artificially turn them into the parallel actions, the beginning of which coincides with the point 0 of the time axis (Fig.4).

This figure contains a set of N parallel actions of an individual in the form of the segments of line, each of which has a random length or duration θ.

In addition the number of segments of lines crossing an arbitrary section θ represents the number of unfinished actions $n(\theta)$ of the individual at the relative moment of time θ. It is also clear that $n(\theta)$ is a decreasing function of time. The number of completed actions by the time θ can be defined as $m(\theta) = N - n(\theta)$, which is an increasing function of time.

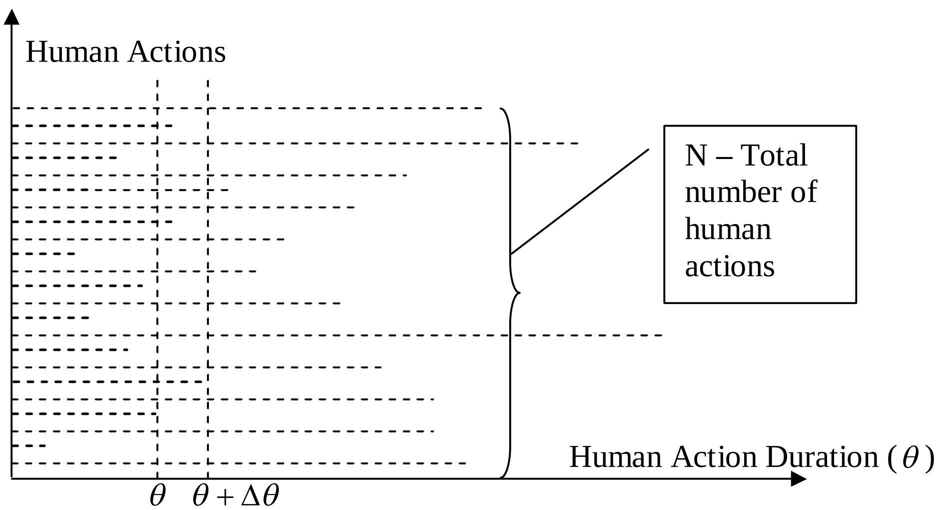


Fig.4 Parallel representation of the serial human actions

This means that the statistical probability $Q^*(\theta)$ of the completion of an arbitrary human action within θ can be defined as

$$Q^*(\theta) = \frac{m(\theta)}{N} = \frac{N - n(\theta)}{N} \qquad (1)$$

Similarly, the probability of incompleteness of a human action can be determined by the formula

$$P^*(\theta) = \frac{n(\theta)}{N} \qquad (2)$$

From these values the probability $Q(\theta)$ represents itself the distribution function of the duration of human actions (Fig.5).

The probability density function of the duration of a single human action is equal to

$$f(\theta) = \frac{dQ}{d\theta}\ . \qquad (3)$$

In its turn the empirical density function of the human action duration $f^*(\theta)$ can be defined as

$$f^*(\theta) = \frac{\Delta n(\theta)}{N\Delta\theta} = \frac{-[n(\theta+\Delta\theta) - n(\theta)]}{N\Delta\theta} \quad . \qquad (4)$$

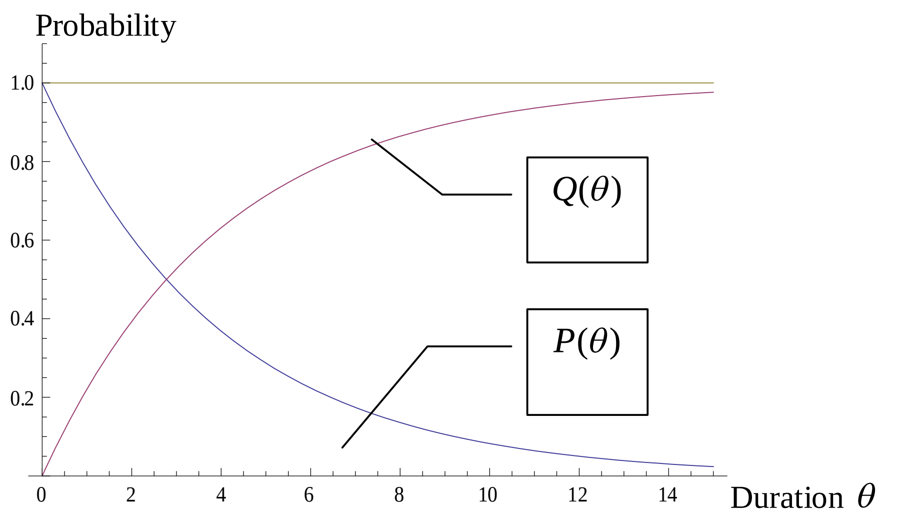


Fig.5 Probabilistic characteristics of the random human action duration

Here $\Delta n(\theta)$ - is the number of completed human actions in the vicinity $\Delta\theta$ of θ.
Histogram of this density function is presented in Fig.6.

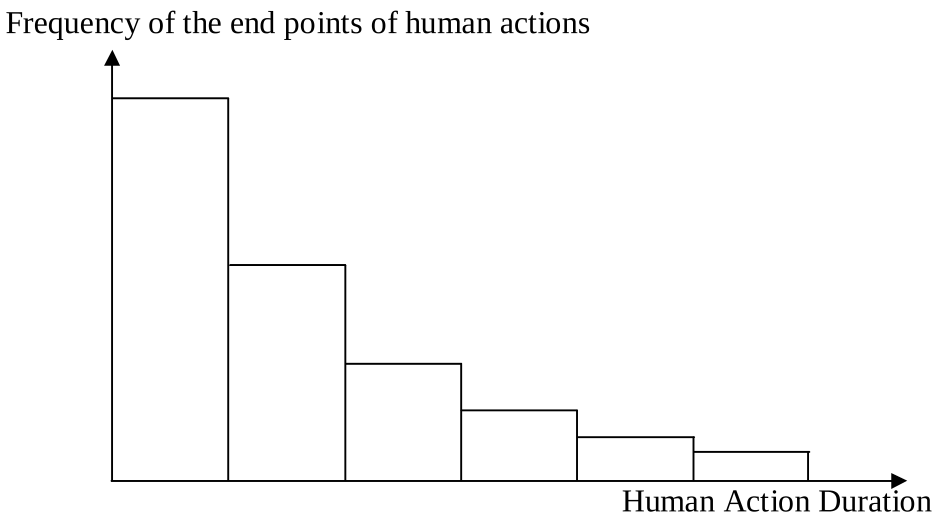


Fig.6. Histogram of the end points of human actions

Probabilistic characteristics of the duration of human actions can be easily determined through the density function $f(\theta)$.

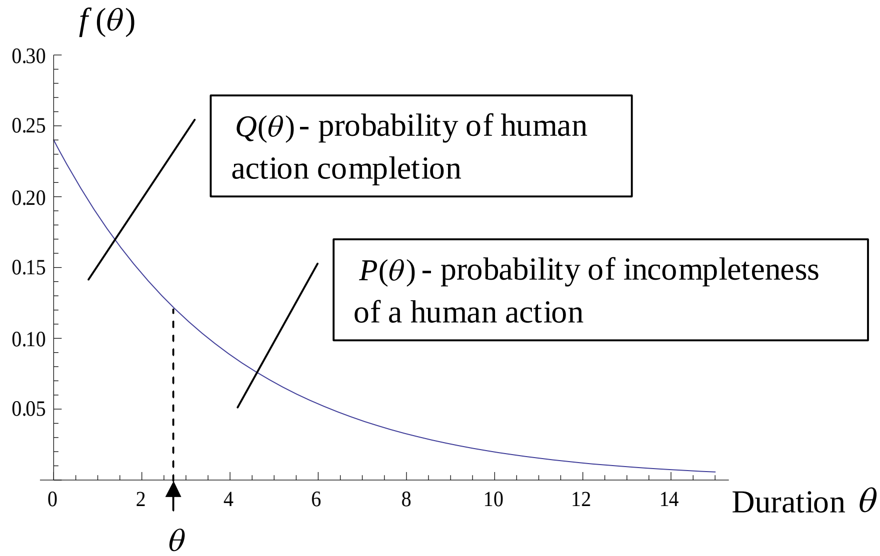


Fig.7 Human action duration distribution density function

Thus, the completion probability of the human single action can be defined as is follows

$$Q(\theta) = \int_0^\theta f(\theta) d\theta . \qquad (5)$$

The probability of incompleteness of a single human action will be

$$P(\theta) = 1 - Q(\theta) = \int_\theta^\infty f(\theta) d\theta \ . \qquad (6)$$

The next value that can be used for the quantitative description of human actions is the intensity of the completion of human actions. Empirically, this value can be defined as

$$\gamma^*(\theta) = \frac{\Delta n(\theta)}{n(\theta)\Delta\theta} = \frac{-\left[n(\theta+\Delta\theta) - n(\theta)\right]}{n(\theta)\Delta\theta} . \qquad (7)$$

According to this definition, the value $\gamma(\theta)\Delta\theta$ is the probability of the completion one of the actions within the duration interval $\Delta\theta$.

Note that the function $\gamma(\theta)$ is not a probability density function of the random variable. This function is not normalized and therefore

$$\int_0^\infty \gamma(\theta) d\theta = \infty .$$

The dynamic model of human actions

Consider the dynamics of the number of uncompleted human actions $n(\theta)$ as a function of θ. In order to do that we need to find out the number of completed actions $\Delta n(\theta)$ in the interval $\Delta\theta$. This value can be defined as

$$\Delta n = -\left[n(\theta + \Delta\theta) - n(\theta)\right]. \tag{8}$$

The minus sign appears because of decreasing character of the function $n(\theta)$.
The same quantity $\Delta n(\theta)$ according to the definition of the function $\gamma(\theta)$ can be defined as is follows

$$\Delta n = \left[n(\theta + \Delta\theta) - n(\theta)\right] = n(\theta)\gamma(\theta)\Delta\theta\,. \tag{9}$$

From here one can obtain the following differential equation

$$\frac{dn}{d\theta} = -\gamma(\theta)n(\theta)\,. \tag{10}$$

Integrating this equation with the initial condition $n(\theta = 0) = N$ will yield

$$n(\theta) = Ne^{-\int_0^\theta \gamma(t)dt}\,. \tag{11}$$

Taking into account expressions (1) and (2) for the probability of incompleteness of a single human action one can obtain

$$P(\theta) = e^{-\int_0^\theta \gamma(t)dt}\,. \tag{12}$$

Probability of the single action completion is

$$Q(\theta) = 1 - P(\theta) = 1 - e^{-\int_0^\theta \gamma(t)dt}\,. \tag{13}$$

The probability density function of the duration of human actions will be

$$f(\theta) = \frac{dQ}{d\theta} = -\frac{dP}{d\theta} = \gamma(\theta)e^{-\int_0^\theta \gamma(t)dt}\,. \tag{14}$$

Average duration of the single human action can be defined by the formula

$$T = \int_0^\infty \theta f(\theta)d\theta\,. \tag{15}$$

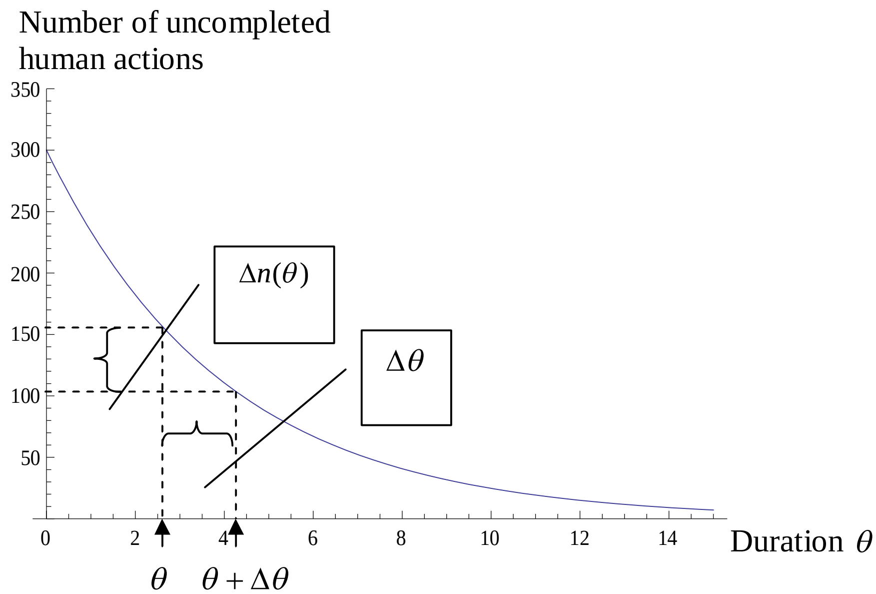


Fig.8 Derivation of the equation for the number of uncompleted human actions

Taking into account that

$$f(\theta) = -\frac{dP}{d\theta} = -P^{'}(\theta) \;, \tag{16}$$

and substituting it into the expression (15) one can obtain

$$T = -\int_0^\infty \theta P^{'}(\theta) d\theta \,. \tag{17}$$

Performing integration by parts from this expression one can obtain

$$T = \int_0^\infty P(\theta) d\theta \;. \tag{18}$$

Substituting into this expression the formula for determining the probability $P(\theta)$ from (12), one can obtain

$$T = \int_0^\infty e^{-\int_0^\theta \gamma(t)dt} d\theta \,. \tag{19}$$

For the probabilistic analysis of the sequence of human actions and for the subsequent analysis of the scheduling risk it is important to obtain an expression for the variance of the duration of a single human action

$$\sigma^2 = \int_0^\infty \theta^2 f(\theta)d\theta - \left(\int_0^\infty \theta f(\theta)d\theta \right)^2 . \qquad (20)$$

Performing the needed transformations and integration, we obtain

$$\sigma^2 = 2\int_0^\infty \theta P(\theta)d\theta - \left(\int_0^\infty P(\theta)d\theta \right)^2 . \qquad (21)$$

Substituting from (12) the value of probability P into this expression one can obtain

$$\sigma^2 = 2\int_0^\infty \theta e^{-\int_0^\theta \gamma(t)dt} d\theta - \left(\int_0^\infty e^{-\int_0^\theta \gamma(t)dt} d\theta \right)^2 . \qquad (22)$$

Thus, all the characteristics associated with the probabilistic representation of a single human action $Q(\theta)$, $P(\theta)$, $f(\theta)$, T and σ^2 depend on the same parameter $\gamma(\theta)$ - intensity of the completion of human actions.

Differential equations for the probabilities $P(\theta)$ and $Q(\theta)$

Using differential equation (10) and expression (2) one can find an equation for the probability of incompleteness of a single human action

$$\frac{dP(\theta)}{d\theta} = -\gamma(\theta)P(\theta) \qquad (23)$$

Taking into account $Q(\theta) = 1 - P(\theta)$, we can have another equation for the probability of completeness of a single human action.

$$\frac{dQ(\theta)}{d\theta} = \gamma(\theta)(1 - Q(\theta)) . \qquad (24)$$

Analysis of the different options of risk functions

The study of the obtained differential equations indicates that they are a convenient means for the development of the new methods for assessing the scheduling risk in the area of project management.

Depending on the complexity of projects and skills of developers the characteristics of the risk functions may vary within a wide range.

The same applies also to the analysis of the timing risks of the single human actions.
The fact that the resulting differential equations (23) and (24) for the analysis of such actions contain only the intensity function $\gamma(\theta)$**,** indicates that by selecting this function one can generate different scheduling risk functions.

For this purpose let's examine some specific forms of the intensity functions $\gamma(\theta)$. The goal of this analysis is to show the mechanisms of the formation of different risk functions with finite and infinite variances (with and without fat tails).

1. **Exponential distribution of the duration of human actions (** $\gamma(\theta) = \gamma = Cons\tan t$ **)**

In this particular case, substituting $\gamma(\theta) = \gamma = Cons\tan t$ into the expression for the density function (14), one can have

$$f(\theta) = \frac{dQ}{d\theta} = -\frac{dP}{d\theta} = \gamma e^{-\gamma\theta} \quad . \qquad (23)$$

Using the expression (23) one can obtain a formula for determining the average duration of a human action

$$T = \int_0^\infty e^{-\gamma\theta} d\theta = \frac{1}{\gamma} . \qquad (24)$$

Calculation of the variance of a single human action duration yields

$$\sigma^2 = 2\int_0^\infty \theta e^{-\gamma\theta} d\theta - \left(\int_0^\infty e^{-\gamma\theta} d\theta \right)^2 = \frac{1}{\gamma^2} = T^2 . \qquad (25)$$

Another important property of the exponential distribution is that if a human action is initiated at the moment of time "0" and is not yet complete at the time of "θ", the distribution of the remaining duration of the human action will be the same as in the instant "0". This property is a direct consequence of the exponential distribution and at first glance seems paradoxical. Other distributions, such as Weibull or Pareto distribution, do not have such properties and are reflecting the dynamics of human actions more adequately.

In real life, the sequences of human actions have complex properties that need to be described by such distribution functions, which are able to adequately describe the delay in work and other problems associated with scheduling risks.

2. Weibull distribution of the duration of human actions

For the analysis we use two-parameter Weibull distribution with distribution function [5, 6]

$$Q(\theta) = 1 - e^{-\gamma\theta^k}, \tag{26}$$

From here for the distribution density function one can obtain

$$f(\theta) = \frac{dQ}{d\theta} = -\frac{dP}{d\theta} = \gamma k \theta^{k-1} e^{-\gamma\theta^k}. \tag{27}$$

Probability of incompleteness of a single human action can be defined as

$$P(\theta) = 1 - Q(\theta) = e^{-\gamma\theta^k}. \tag{28}$$

The average duration of human actions can be estimated by the formula

$$T = \int_0^\infty e^{-\gamma\theta^k} d\theta = \gamma^{-\frac{1}{k}} \Gamma(1 + \frac{1}{k}). \tag{29}$$

Here $\Gamma(x)$ is the gamma function.
The variance of the single human action can be defined as

$$\sigma^2 = \gamma^{-\frac{2}{k}} \left[\Gamma\left(1 + \frac{2}{k}\right) - \Gamma^2\left(1 + \frac{2}{k}\right) \right]. \tag{30}$$

The intensity function of the human action completion has the form

$$\gamma(\theta) = \frac{f(\theta)}{P(\theta)} = \gamma k \theta^{k-1}. \tag{31}$$

Changing the parameter k in Weibull distribution allows making a smooth transition from distributions with finite variance to the distributions with heavy tail.

Fig.9 shows a smooth transition, where the increasing parameter k has four discrete values, thereby generating four different distributions.

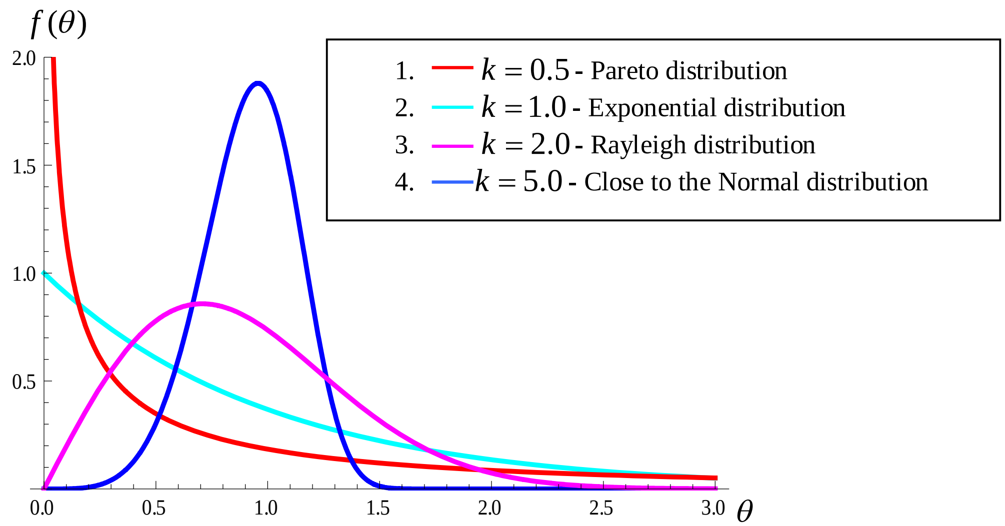


Fig.9 Different presentations of the same Weibull distribution

In this regard, the parameter k may serve as a bridge between the degree of difficulty of human actions and the distribution of the duration of human actions. In quantitative terms, we may assume that as a first approximation, this parameter is proportional to the difference $D_w - D_{av}$.

$$k = \beta(D_w - D_{av}), \qquad (32)$$

where β - coefficient of proportionality.
Substituting the expression (32) into the Weibull distribution density function (27) we can have

$$f(\theta) = \gamma\beta(D_w - D_{av})\theta^{\beta(D_w - D_{av})-1}e^{-\gamma\theta^{\beta(D_w - D_{av})}}. \qquad (33)$$

This relationship establishes a direct functional link between the risk function and the parameters of the action difficulty D_{av} and the human capacity D_w to overcome that difficulty.

The above results indicate that the risk function cannot be chosen arbitrary, but must be conditioned by the parameters of the work process and by the parameters of the people themselves, as nonlinear converters of the action's difficulty into the action's duration.

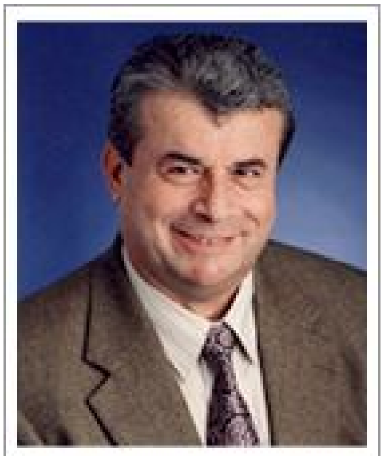
Conclusions

- For the solution of the problem of human action duration estimation it is necessary to consider the human beings as nonlinear converters of the random difficulties of actions into their durations,
- Probabilistic characteristics of the duration of human actions can be obtained by the processing statistical data,
- Introduction of the notion of intensity of the completion of human actions allows deriving a differential equation with regard to the probabilistic parameters of the duration human actions,
- Human action duration distribution function can be defined using action completion intensity functions only,
- Different human action completion intensity functions can generate different distribution functions for human action duration,
- Depending on circumstances the same mathematical model for human action duration can generate distributions with and without fat tail,
- Scheduling risk function cannot be chosen arbitrary; instead they can be derived analytically.

References

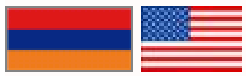
1. Pavel Barseghyan. (2009). Principles of Top-Down Quantitative Analysis of Projects. Part 1: State Equation of Projects and Project Change Analysis. PM World Today – May 2009 (Vol XI, Issue **V**). 16 pages.
2. Pavel Barseghyan (2010) “Parkinson’s Law, Overtime Work and Human Productivity”. *PM World Today* – February 2010 (Vol XII, Issue II). 11 pages.
3. Pavel Barseghyan (2009) “Human Effort Dynamics and Schedule Risk Analysis”. *PM World Today* – March 2009 (Vol XI, Issue III). 13 pages.
4. (Pavel Barseghyan. (2009). Task Assignment as a Crucial Factor for Project Success (Probabilistic Analysis of Task Assignment) ”. *PM World Today* – April 2009 (Vol XI, Issue I**V**). 15 pages.
5. B. Gnedenko, (2005), The Theory of Probability (AMS Chelsea Publishing).
6. Kapur, K.C., and Lamberson, L.R., (1977), *Reliability in Engineering Design*, John Wiley & Sons, New York.

About the Author



Pavel Barseghyan, PhD

Author



Dr. Pavel Barseghyan is a consultant in the field of quantitative project management, project data mining and organizational science. He is the founder of Systemic PM, LLC, a project management company. Has over 40 years experience in academia, the electronics industry, the EDA industry and Project Management Research and tools development. During the period of 1999-2010 he was the Vice President of Research for Numetrics Management Systems. Prior to joining Numetrics, Dr. Barseghyan worked as an R&D manager at Infinite Technology Corp. in Texas. He was also a founder and the president of an EDA start-up company, DAN Technologies, Ltd. that focused on high-level chip design planning and RTL structural floor planning technologies. Before joining ITC, Dr. Barseghyan was head of the Electronic Design and CAD department at the State Engineering University of Armenia, focusing on development of the Theory of Massively Interconnected Systems and its applications to electronic design. During the period of 1975-1990, he was also a member of the University Educational Policy Commission for Electronic Design and CAD Direction in the Higher Education Ministry of the former USSR. Earlier in his career he was a senior researcher in Yerevan Research and Development Institute of Mathematical Machines (Armenia). He is an author of nine monographs and textbooks and more than 100 scientific articles in the area of quantitative project management, mathematical theory of human work, electronic design and EDA methodologies, and tools development. More than 10 Ph.D. degrees have been awarded under his supervision. Dr. Barseghyan holds an MS in Electrical Engineering (1967) and Ph.D. (1972) and Doctor of Technical Sciences (1990) in Computer Engineering from Yerevan Polytechnic Institute (Armenia). Pavel can be contacted at pavel@systemicpm.com.