CHEMICAL ENGINEERING TRANSACTIONS
VOL. 33, 2013
A publication of AIDIC
The Italian Association of Chemical Engineering
Online at: www.aidic.it/cet
Guest Editors: Enrico Zio, Piero Baraldi

ISBN 978-88-95608-24-2; **ISSN** 1974-9791
DOI: 10.3303/CET1333031

Quantifying Uncertainty for System Failure Prognostics: Enhanced Approach Based on Interval-Valued Probabilities

Victor G. Krymsky

Ufa State University of Economics and Service, 145 Chernyshevsky Street, Ufa, 450078, Russia
vikrymsky@mail.ru

Failure prognostics being performed for any complex industrial system typically meets serious obstacles caused by uncertainty. This uncertainty actually has a hierarchical structure due to the peculiarities of its influence to the prognostics results. Namely, to quantify the chances of a system failure we normally have to apply probabilistic (non-deterministic) model. Such a model reflects the randomness and the diversity of the factors initiating a failure (the 'first order uncertainty'). However the estimated probability of the failure may differ from its real value by the so-called 'bias'. So the imperfectness of the probabilistic models becomes the source of the 'second order uncertainty'. This means that the interval-valued probabilities can often describe the reality more adequately than single valued ones.
The paper presents the novel approach to system life prognostics based on the usage of the interval-valued statistical characteristics. The problem statement remarkably has many features similar to those considered in the classical publications on the theory of imprecise probabilities. Nevertheless the methodology used to find the prognostics result is very different and involves the technique from the optimal control theory. It relies on the state space representation of the system life cycle and Pontryagin's principle of maximum modified to optimization problems with the integral and non-integral types of constraints. This novelty allows generalizing the ways to process additional information on the system reliability efficiently and makes the resulted intervals for the probability values tighter. In turn the decrease of the corresponding imprecision provides favorable conditions for justified decisions on the failure prevention, e.g. maintenance planning.

1. First and second order uncertainties and system failure prognostics

The probabilistic models have become the classical technique for treating uncertainties. To apply this technique to most practical problems we need precise information on the distributions of random factors and on the values of probabilities characterizing initiating events. When analyzing the possibilities of a system failure or industrial accident one normally deals with probabilities of finding the system in certain states and conditional probabilities of the system parameters changes due to the change in the system state. If some of the initial probabilities are unknown or imperfectly known we cannot expect to obtain an accurate result (e.g. probability of the system failure). Let us illustrate by an example.
Consider gas transmission pipeline – an industrial system which proneness to failure (rupture) is influenced by a number of factors (third parties actions, corrosion, construction defects, land movement, etc.). Different combinations of these discretized factors will therefore produce specific states of the system. Each state θ_i, $i = 1, 2, ..., n$, occurs with a probability $P(\theta_i)$ and is associated with the definite values of some parameter X important for the failure identification. In the case of the gas pipeline one may think of the pressure drop in the pipe (Vianello and Maschio, 2011).
Assume the range of X values can be also discretized into the sequence of the ordered discrete values x_j, $j = 1, 2, ..., m$, and each x_j is associated with a probability $P(x_j / \theta_i)$ given the condition of finding the system in the state θ_i. Denote by x^* the critical value of X such that the system failure is recognized as identified if X exceeds x^*. The failure probability can then be obtained directly from the total probability formula:

Please cite this article as: Krymsky V., 2013, Quantifying uncertainty for system failure prognostics: enhanced approach based on interval-valued probabilities, Chemical Engineering Transactions, 33, 181-186 DOI: 10.3303/CET1333031

$$P_f = \Pr(X \geq x^*) = \sum_{\forall j: x_j \geq x^*} \sum_{i=1}^{n} P(x_j / \theta_i) \cdot P(\theta_i). \tag{1}$$

The expression (1) emphasizes that failure prognostics has to be done in the conditions of uncertainty, and involves consideration of random events (e.g. different states occurrences) and stochastic values of the parameter X. We may call it the 'first order uncertainty'.

At the same time the practical experience shows that real value P_f^R of the failure probability differs from P_f due to imprecision and incompleteness of initial statistical data which influence the accuracy of $P(\theta_i)$ and $P(x_j / \theta_i)$. In fact $P_f^R = P_f + E^*$, where $E^* \in [E_{\min}^*, E_{\max}^*]$ is the so-called 'bias' (Zio and Apostolakis, 1996). This bias reflects the 'second order uncertainty' and implies that instead of the point value for the failure probability (classical approach) we have to deal with P_f^R belonging to the interval

$$P_f^R \in [P_f + E_{\min}^*, P_f + E_{\max}^*]. \tag{2}$$

The expression (2) suggests that the interval-valued probabilities can provide better representation of the hierarchical uncertainty structure combining first and second order uncertainties than the point-valued ones. However the existing algorithms exploiting interval-valued probabilities have a number of disadvantages which become a 'stumbling block' for reliability applications (Kozine and Krymsky, 2009). We analyse the features of the approach based on the interval-valued probabilities and demonstrate the way in which it can be improved for practical needs.

2. Interval-valued (imprecise) probabilities: typical problem statements

Imprecise prevision theory (IPT) started by fundamental publications of Walley (1991) and Kuznetsov (1991) deals with exactly the uncertainty models containing interval-valued statistical characteristics. An important advantage of IPT is its capability to combine both statistical data and experts judgments when estimating the lower and the upper bounds on probabilities and other relevant characteristics. Such estimates can be obtained by solving problems of linear programming type without introducing additional assumptions on a probability distribution.
Traditional IPT problem statement in one-dimensional case considers the following constraints:

$$\rho(t) \geq 0, \tag{3}$$

$$\int_0^{T^*} \rho(t)dt = 1, \ \underline{a_i} \leq \int_0^{T^*} f_i(t)\rho(t)dt \leq \overline{a_i}, \ i = 1, 2, ..., n. \tag{4}$$

Here $\rho(t)$ is unknown probability density function (pdf) of the random variable T, outcomes of which belong to the interval $[0, T^*]$, $f_i(t)$, $i = 1, 2, ..., n$, are given real-valued positive functions (so-called 'gambles'), $\underline{a_i}$, $\overline{a_i} \in \Re^+$, $i = 1, 2, ..., n$, are given numbers. In reliability applications T typically means time to failure.
It is necessary to find the coherent lower and upper previsions $\underline{M}(g)$ and $\overline{M}(g)$ for expectation $M(g)$ of some other non-negative function ('gamble') $g(t)$ subject to constraints (3) and (4). In turn

$$\underline{M}(g) = \inf_{\rho(t)} \int_0^{T^*} g(t)\rho(t)dt, \ \overline{M}(g) = \sup_{\rho(t)} \int_0^{T^*} g(t)\rho(t)dt. \tag{5}$$

In some practical applications (e.g. failure prognostics) the bounds of the intervals $[\underline{M}(g), \overline{M}(g)]$ generated by traditional IPT methods, however, are not tight enough, so the resulted uncertainty is too high for providing justified decision support. The additional constraints can be introduced to reduce the width of the prevision intervals.

The first enhancement of the approach exploits an idea of restricting ourselves to the case in which the unknown pdf values are reasonably bounded by the definite positive number K elicited from the expert judgments (Krymsky, 2006). Actually this requirement takes the form of inequality

$$\rho(t) \le K\,,\ KT^{*} \le 1. \tag{6}$$

The inequality (6) allows excluding unbounded pdf's (e.g. linear combinations of Dirac's δ-functions) from the set of densities $\rho(x)$ for which the expectation $M(g)$ attains its maximum or minimum. One can also require the derivatives of $\rho(x)$ to be bounded (Kozine and Krymsky, 2009). Both bounded pdf's and their derivatives give the possibility to make the bounds of the interval $[\underline{M}(g), \overline{M}(g)]$ much tighter.

The last enhancement is related to the problem statement which deals with the bounded failure rates (Kozine and Krymsky, 2012) and assumes that

$$\underline{\lambda} \le \lambda(t) \le \overline{\lambda}, \tag{7}$$

where $\lambda(t) = \rho(t)/\left(1 - F(t)\right)$ is the failure rate, $\underline{\lambda},\ \overline{\lambda} \in \Re^{+}$ are given numbers, $F(t)$ is cumulative distribution function (cdf):

$$F(t) = \int_{0}^{t} \rho(t)dt. \tag{8}$$

Thanks to the bounded failure rates, one can extend the interval of possible time to failure values till infinity ($T^{*} \to \infty$). Such a concept is significantly more adequate for representing the real situations as actually nobody a priori knows how long the system can work before the failure occurs.

Note that the constraints (6) or (7) do not allow solving the corresponding optimization problems in the framework of the traditional IPT methodology based on linear programming algorithms (Walley, 1991; Kuznetsov, 1991). The crux is that we cannot convert the primal problem to a dual form as typically done when linear programming technique is applied. Namely, in this case the dual problem will contain an infinite number of variables. The appeared difficulty has been overcome by the application of variational calculus methods (Krymsky, 2006; Kozine and Krymsky, 2009, 2012).

The obtained results confirm that the accuracy of reliability interval assessments has been improved if the constraints (6) or (7) are utilized. However the approach has not been sufficiently generalized yet. In particular, it cannot operate with constraints of the types (6) and (7) simultaneously. In this paper we continue development of the approach taking into account that the models under investigation are very similar to those considered within optimal control theory (Pontryagin et al, 1962; Hestenes, 1966).

3. Interval probabilities and optimal control model

The control models are used to study the system dynamics which is usually described by the differential equations in a state space form (Brogan, 1991):

$$dx_l(t)/dt = \varphi_l\left(x_1(t), x_2(t), ..., x_r(t), u(t), t\right),\ l = 1, 2, ..., r. \tag{9}$$

Here $u(t)$ is control (input) signal, $x_l(t)$, $\varphi_l(\bullet)$, $l = 1, 2, ..., r$, are state space coordinates and given functions respectively. In failure prognostics one has to consider two functions which interrelations correspond to the differential equation – pdf and cdf. Denote $u(t) = \rho(t)$, $x_1(t) = F(t)$. Then

$$dx_1(t)/dt = \varphi_1\left(x_1(t), u(t), t\right) = u(t). \tag{10}$$

Each curve in the Cartesian plane (Figure 1) reflects a possible relation between $x_1(t)$ and $u(t)$ when time t changes from zero to T^{*}. These curves are graphical trajectories of a system life cycle. A particular curve starts at some point on Y-axis when $t = 0$ and continues in the 'left-to-right' direction (as the derivative $u(t)$ of $x_1(t)$ is non-negative only if $x_1(t)$ is a non-decreasing function). When $t \to T^{*}$ the

trajectory attains its terminal point located at the vertical line for which $x_1(T^*)=F(T^*)=1$ (Figure 1). The situation in which $T^* \to \infty$ is also allowed as the so-called 'infinite horizon control' (Carlson et al, 1991).
Most types of constraints can be visualized in the same plane. For instance, the bounded pdf (6) requires the existence of a horizontal line $u(t)=K$ (Figure 2). The bounds on the failure rates introduced by the inequalities (7) are depicted by the straight lines described by the linear algebraic equations $u(t)=\underline{\lambda}(1-x_1(t))$ and $u(t)=\overline{\lambda}(1-x_1(t))$. As the result only the trajectories lying inside the area bounded by the aforementioned straight lines (Figure 2) satisfy the constraints (6) and (7).

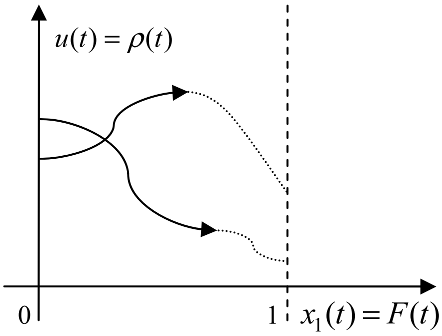

Figure 1: trajectories of the system life cycle in the Cartesian plane

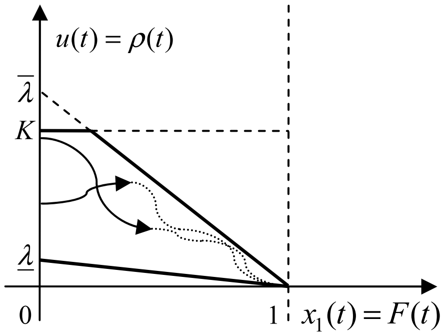

Figure *2: trajectories of the system life cycle in case of bounded pdf and failure rate*

We aim at seeking the control functions $u(t)$ and the corresponding system life cycle trajectories for which the expectation $M(g)$ attains its maximum (minimum) subject to constraints (3), (4), (6), (7). This optimization problem includes various types of constraints. In particular, expressions (4) represent so-called 'isoperimetric' constraints (they set the bounds on integrals depending on $u(t)=\rho(t)$). On the contrary (3), (6) and (7) belong to the type of 'holonomic' (non-integral) constraints.
The solution of the formulated optimal control problem can be obtained via application of the principle of maximum (Pontryagin et al, 1962). Its modification to the case of combined isoperimetric and holonomic constraints was developed by Hestenes (1966). We introduce the following notation:

$$\begin{aligned}
&\varphi_0(x_1(t),u(t),\mathbf{b},t)=g(t)u(t);\\
&h_k(t)=f_k(t)u(t) \text{ if } k=1,2,...,n \text{ or } h_k(t)=-f_{k-n}(t)u(t) \text{ if } k=n+1,n+2,...,2n;\\
&h_{2n+1}(t)=u(t);\ S_k(\mathbf{b})=-\underline{a}_k \text{ if } k=1,2,...,n \text{ or } S_k(\mathbf{b})=\overline{a}_k \text{ if } k=n+1,n+2,...,2n;\\
&\xi_1(x_1(t),u(t),\mathbf{b},t)=u(t);\ \xi_2(x_1(t),u(t),\mathbf{b},t)=K-u(t);\\
&\xi_3(x_1(t),u(t),\mathbf{b},t)=u(t)/(1-x_1(t))-\underline{\lambda};\ \xi_4(x_1(t),u(t),\mathbf{b},t)=\overline{\lambda}-u(t)/(1-x_1(t)),
\end{aligned} \tag{11}$$

where $\mathbf{b}=(b_1,b_2,...,b_{2n})$ is the vector of control parameters.
Assume the optimal solution of the formulated problem is $u(t)=u^*(t)$, $x_1(t)=x_1^*(t)$, $\mathbf{b}=\mathbf{b}^*$.
Then according to the results of optimal control theory (Hestenes, 1966) there exist the functions

$$\begin{aligned}
&H(x_1(t),u(t),\mathbf{b},t,\psi_1(t),\boldsymbol{\mu})\\
&=\psi_0\varphi_0(x_1(t),u(t),\mathbf{b},t)+\psi_1(t)\varphi_1(x_1(t),u(t),\mathbf{b},t)+\sum_{k=1}^{2n+1}\mu_k h_k(x_1(t),u(t),\mathbf{b},t),
\end{aligned} \tag{12}$$

$$L(x_1(t),u(t),\mathbf{b},t,\psi_1(t),\boldsymbol{\mu})=H(x_1(t),u(t),\mathbf{b},t,\psi_1(t),\boldsymbol{\mu})+\sum_{j=1}^{4}\nu_j(t)\xi_j(x_1(t),u(t),\mathbf{b},t),. \tag{13}$$

in which $\boldsymbol{\mu}=(\mu_1,\mu_2,...,\mu_{2n+1})$ is the vector of multipliers,

such that the following relations hold:
(a) The multipliers ψ_0, μ_k, $k = 1, 2, ..., 2n+1$, are constant; $\psi_0 \geq 0$ and $\mu_k \geq 0, \quad k = 1, 2, ..., 2n+1$, with

$$\mu_k \left(\int_0^{T^*} h_k \left(x_1^*(t), u^*(t), \mathbf{b}^*, t \right) dt + S_k(\mathbf{b}^*) \right) = 0 \,, \; k = 1, 2, ..., 2n+1 \,. \tag{14}$$

(b) The multipliers $\nu_j(t)$, $j = 1, 2, 3, 4$, are piecewise continuous and are continuous over each interval of continuity of $u^*(t)$. Moreover for each $j = 1, 2, 3, 4$ we have

$$\nu_j(t) \geq 0 \,, \; \nu_j(t) \xi_j \left(x_1^*(t), u^*(t), \mathbf{b}^*, t \right) = 0 \,. \tag{15}$$

(c) The multiplier $\psi_1(t)$ is continuous and satisfies the adjoint equation

$$-d\psi_1(t)/dt = (\partial/\partial x_1) H\left(x_1(t), u^*(t), \mathbf{b}^*, t, \psi_1(t), \boldsymbol{\mu} \right) \Big|_{x_1(t) = x_1^*(t)}. \tag{16}$$

Note that in our case $d\psi_1(t)/dt = 0$, so $\psi_1(t) = \psi_1 = const$.
The maximum principle is stated by the inequality

$$H\left(x_1^*(t), u^*(t), \mathbf{b}^*, t, \psi_1(t), \boldsymbol{\mu} \right) \geq H\left(x_1^*(t), u(t), \mathbf{b}^*, t, \psi_1(t), \boldsymbol{\mu} \right) \tag{17}$$

which holds for all $[x_1^*(t), u(t), \mathbf{b}^*, t] \in A$, where A is the set of acceptable solutions.
The maximum principle implies that

$$(\partial/\partial u) L\left(x_1^*(t), u(t), \mathbf{b}^*, t, \psi_1(t), \boldsymbol{\mu} \right) \Big|_{u(t) = u^*(t)} = 0. \tag{18}$$

Analyzing equation (18) subject to (12) – (16) one can conclude that the optimal pdf consists of some intervals of continuity, and at each interval either pdf equals K or the failure rate $\lambda(t)$ equals $\underline{\lambda}$ or $\overline{\lambda}$.

4. Example

Assume we are interested in knowing bounds $\underline{\tau}$ and $\overline{\tau}$ on the mean time to failure

$$\tau = \int_0^\infty t\rho(t)dt = \int_0^\infty (1 - F(t))dt \tag{19}$$

of a system and the following constraints are known: $\underline{\lambda} \leq \lambda(t)$, $\rho(t) \leq K$, $\underline{\lambda} < K$ and

$$R(q) = 1 - \int_0^\infty I_{[0,q]}(t)\rho(t)dt = p. \tag{20}$$

Here $R(q)$ is system reliability at time q, $I_{[0,q]}(t)$ equals 1 if $t \in [0,q]$ or equals 0 otherwise.

The pdf $u^{*(1)}(t) = \rho(t)$ for which τ attains its minimum $\underline{\tau}$ contains two intervals of continuity: $u^{*(1)}(t) = K$ if $t \leq t_1$ and $u^{*(1)}(t) = (1 - Kt_1)\underline{\lambda} \exp\left(-\underline{\lambda}(t - t_1)\right)$ if $t > t_1$. This yields $\underline{\tau} = Kt_1^2/2 + (1 - Kt_1)(t_1 + 1/\underline{\lambda})$. In turn t_1 can be found as the root of the equation $(1 - Kt_1)\exp\left(-\underline{\lambda}(q - t_1)\right) = p$.

The pdf $u^{*(2)}(t) = \rho(t)$ for which τ attains its maximum $\overline{\tau}$ contains three intervals of continuity: $u^{*(2)}(t) = \underline{\lambda}\exp(-\underline{\lambda}t)$ if $t \leq t_2$, $u^{*(2)}(t) = K$ if $t_2 < t \leq q$ and $u^{*(2)}(t) = p\underline{\lambda}\exp\left(-\underline{\lambda}(t - q)\right)$ if $t > q$. This leads to the result $\overline{\tau} = 1/\underline{\lambda} - (1/\underline{\lambda} + t_2)\exp(-\underline{\lambda} t_2) + K(q^2 - t_2^2)/2 + p(q + 1/\underline{\lambda})$. In turn t_2 can be found as the root of the equation $(1 - K(q - t_2))\exp(-\underline{\lambda} t_2) = p$.

Now we compare the obtained results with the solutions to the corresponding traditional IPT problem in which we impose no restrictions for the upper bound on pdf and the lower bound on failure rate. In fact, a priori information can be incorporated in traditional problem statement only in the form of Equation (20). Applying the methods of linear programming to the discussed traditional problem one can estimate the lower bound on mean time to failure as $\underline{\tau}_{trad} = qp$.

Suppose $\underline{\lambda} = 0.0001\,h^{-1}$, $K = 0.0008\,h^{-1}$, $q = 700\,h$ and $p = 0.9$, then $\underline{\tau}_{trad} = 630.0\,h$, $t_1 = 46.3\,h$, $\underline{\tau} = 9{,}674.9\,h$ and $\Delta(\underline{\tau}) = \underline{\tau} - \underline{\tau}_{trad} = 9{,}044.9\,h$. The obtained difference indicates that the lower bound on mean time to failure moves to the right when additional constraints are utilized. In turn the difference $\Delta(\overline{\tau}) = \overline{\tau}_{trad} - \overline{\tau}$, where $\overline{\tau} = 9{,}697.8\,h$, will be arbitrarily large as $\overline{\tau}_{trad} \to \infty$ yielding more accurate estimate for the upper bound on mean time to failure. Hence the bounds of the resulting interval $[\underline{\tau}, \overline{\tau}]$ become significantly tighter than those of the interval $[\underline{\tau}_{trad}, \overline{\tau}_{trad}]$.

5. Conclusions

Interval-valued probabilities are applied to quantifying the uncertainty within the failure prognostic procedures as they adequately reflect the lack of knowledge in the initial data. However the resulting intervals for the statistical characteristics of the system are often too wide for obtaining an accurate failure prognosis. This drawback can be overcome by introducing additional constraints (e.g. bounded probability densities, bounded failure rates) mostly based on expert judgments. The crux is that such an approach does not allow solving the corresponding problems by the linear programming methods as it has been done in the traditional theoretical studies. In this paper we have proposed and developed an alternative approach that exploits the methods of optimal control theory, and demonstrated its applicability.

Acknowledgments

The author expresses sincere gratitude to Dr. Igor Kozine (Technical University of Denmark) for cooperation created the basis for in-depth study of the interval-valued probabilistic uncertainty models.

References

Brogan W.L., 1991, Modern Control Theory. Prentice Hall, Englewood Cliffs, NJ, the USA.

Carlson D.A., Haurie A.B., Leizarowitz A., 1991, Infinite Horizon Optimal Control. Deterministic and Stochastic Systems. Springer Verlag, Berlin, Germany.

Hestenes M.R., 1966, Calculus of Variations and Optimal Control Theory. John Wiley & Sons, New York, the USA.

Kozine I.O., Krymsky V.G., 2009, Bounded densities and their derivatives: extension to other domains, Journal of Statistical Theory and Practice, 3, 25-38.

Kozine I.O., Krymsky V.G., 2012, An interval-valued reliability model with bounded failure rates, Int. Journal of General Systems, 41(8), 760-773.

Krymsky V.G., 2006, Computing interval bounds for statistical characteristics under expert-provided bounds on probability density functions, Lecture Notes in Computer Science, 3732, Springer Verlag, 151-160.

Kuznetsov V.P., 1991, Interval Statistical Models. Radio i Svyaz, Moscow (in Russian).

Pontryagin L.S., Boltyanskii V.G., Gamkrelidze R.V., Mishchenko E.F., 1962, The Mathematical Theory of Optimal Processes. Interscience Publishers, New York, the USA.

Vianello C., Maschio G., 2011, Risk analysis of natural gas pipeline: case study of a generic pipeline, Chemical Engineering Transactions, 24, 1309-1314, DOI: 10.3303/CET1124219.

Walley P., 1991, Statistical Reasoning with Imprecise Probabilities. Chapman and Hall, London, UK.

Zio E., Apostolakis G.E., 1996, Two methods for the structured assessment of model uncertainty by experts in performance assessments of radioactive waste repositories, Reliability Engineering and System Safety, 54, 225-241.