Olgierd Hryniewicz
Systems Research Institute, Warsaw, Poland

Confidence bounds for the reliability of a system from subsystem data

Keywords

reliability, system, confidence intervals, binomial data, exponential data

Abstract

The paper is concerned with the construction of lower bounds for the reliability of a system when statistical data comes from independent tests of its elements. The overview of results known from literature and obtained under the assumption that elements in a system are independent is given. It has been demonstrated using a Monte Carlo experiment that in the case when these elements are dependent and when their dependence is described by Clayton and Gumbel copulas these confidence bounds are not satisfactory. New simple bounds have been proposed which in some practical cases have better properties than the classical ones.

1. Introduction.

Reliability indices of complex systems can be estimated from the results of lifetime tests. When a system is treated as one entity we can distinguish two different types of reliability tests. In the first one, we observe consecutive failures of a system, and after each of them a failed system is completely renewed. In such case, random times between consecutive failures are described by independent random variables having and identical probability distributions. If this assumption is true, we can estimate a required reliability characteristic using a sample of observed lifetimes. In the second case, we have to observe several identical systems working in the same conditions. Times to first failures of these systems constitute a sample which may be used for the estimation of the considered reliability characteristic. In both cases, however, we need to have either sufficiently long time of test or sufficiently large number of observed systems. Both these requirements are seldom met in practice. Thus, this method of the reliability estimation is rarely used in practice despite the fact that from a statistical point of view the required estimators are obtained in the simplest possible way. Moreover, in such a case we do not profit from the information about the structure of the considered system, and from the knowledge of times to failure of its elements.

In practice we are frequently faced with a different problem: how to evaluate reliability characteristics of a system on its design stage. There exist many methods for the prediction of reliability using available statistical data. In this paper we consider the simplest one, when we can utilize the results of reliability tests of system's elements performed in presumably the same conditions as the conditions of work of the designed complex system.

Research studies on statistical methods aimed at the estimation of system's reliability using the results of reliability tests of its elements were initiated independently in the 1950s in the United States and the Soviet Union, where they were performed by prominent mathematicians and statisticians. Some strong mathematical results were obtained, and these results can be used for both point and interval estimation of system's reliability using the data obtained for its elements or subsystems. In this paper we will focus our attention on the interval estimation. The reason for the importance of the results of this type stems from practice. Usually we can use scarce reliability data, and thus the obtained point estimators are not very precise. Therefore, we need to know some lower bounds for the predicted reliability characteristics.

Preliminary analysis of the theoretical results shows undoubtedly that even in the cases of simple systems exact analytical methods require utilization of complex mathematical tools such as nonlinear mathematical programming. On the other hand, interesting approximate results, obtained mainly by American researchers, can be used in practice when a sufficiently large number of failures have been observed. For this reasons already in the 1980s the

reliability theoreticians lost their interest in further research in this area. However, the problem is stall interesting for practitioners who need approximate, or even heuristic, methods which may be used for the prediction of reliability using existing statistical data.

The purpose of this paper is two-fold. In first six sections we give a short overview of different methods for the construction of confidence intervals for the reliability of systems. In all these methods it has been assumed that the elements of a system are independent. In the last section of the paper we present new bounds which in certain practical cases are robust to deviations from the assumption of independence.

2. General methodology for the evaluation of system's reliability

Evaluation of reliability of complex systems became the subject of intensive theoretical investigation in the beginning of 1960s. Fundamental results were summarized in the famous book by Barlow and Proschan [1]. In the mathematical models considered in [1] it is assumed that both the system as a whole, and system's elements at any time instant $t>0$ are either in the state of functioning (or failure-free state), when the random variable $X(t)$ describing the reliability state adopts the value 1, or in the state of failure, when this random variable adopts the value 0. When the considered system consists of m elements, then its reliability state is described by the random vector $\mathbf{X}=(X_1,X_2,\ldots X_m)$, and the probability of the observation of any reliability state is given by

$$P(\mathbf{X})=\prod_{i=1}^{m} p_i^{X_i}(1-p_i)^{1-X_i}\,, \tag{1}$$

where

$$p_i = P(X_i=1)=E(X_i),\ i=1,\ldots,n\,. \tag{2}$$

In the above formulae we have omitted time t assuming that in case of specific calculations it adopts the same value for all components of the random vector.

Reliability state of the whole system depends on the states of all individual system's elements. Denote by Ω the set of all 2^m possible states of system's elements. We can divide this set into two exclusive subsets: the subset of all functioning states of the system G, and the subset of all failure states of this system $\overline{G}$ ($G\cup\overline{G}=\Omega$). The function

$$\Psi(\mathbf{X})=\begin{cases}1 & \mathbf{X}\in G\\ 0 & \mathbf{X}\in \overline{G}\end{cases} \tag{3}$$

is called the *structure function*, and it describes the relation between reliability state of the whole system and reliability states of its elements. The effective construction of this function is the subject of numerous research works. Particular results may be found in all classical textbooks on reliability [1], [2].

Probability that the considered system is in the failure-free state depends on the vector $\mathbf{p}=(p_1,p_2,\ldots,p_m)$ that describes the probabilities of failure-free functioning of system's elements, and system's reliability structure function. It is given by the function called the reliability function which is given by the following formula

$$R(\mathbf{p})=P(\mathbf{X}\in G)=E[\Psi(\mathbf{X})]=\sum_{\mathbf{X}\in G}\Psi(\mathbf{X})\prod_{i=1}^{m} p_i^{X_i}(1-p_i)^{1-X_i}\quad . \tag{4}$$

Below, we present the respective formulae for the reliability structures which are most frequently met in practice.

a) In case of a system with series reliability structures which consists of m groups of identical $n_i, i=1,\ldots,m$ elements we have:

$$R(\mathbf{p})=\prod_{i=1}^{m} p_i^{n_i}\,. \tag{5}$$

b) For the system with a parallel reliability structure which consists of m elements the respective formula is given by

$$R(\mathbf{p})=1-\prod_{i=1}^{m}(1-p_i)=1-\prod_{i=1}^{m} q_i\,. \tag{6}$$

c) In case of a series-parallel reliability system which consists of m connected in series groups, where each of these groups consists of n_i connected in parallel identical elements, the reliability function is given by the formula:

$$R(\mathbf{p})=\prod_{i=1}^{m}\left[1-(1-p_i)^{n_i}\right]. \tag{7}$$

d) For a parallel-series system consisting of m connected in parallel groups, where each of these groups consists of n_i identical elements connected in series, the reliability function is given by the formula:

$$R(\mathbf{p}) = 1 - \prod_{i=1}^{m}\left[1 - \prod_{j=1}^{n_i} p_{ij}\right]. \quad (8)$$

In formulae (5) – (8) p_{ij} denotes the probability that the *j*-th element in the *i*-th subsystem is in a failure-free state.

The systems with structures described above belong to a more general class of systems called coherent systems, or systems with monotonic structure. The system has monotonic structure if

$$\Psi(\mathbf{X}) \geq \Psi(\mathbf{Y}) \quad (9)$$

holds when $X_i \geq Y_i, i = 1,\ldots,m$, and when

$$\Psi(\mathbf{0}) = 0, \ldots \Psi(\mathbf{1}) = 1, \quad (10)$$

with **0**=(0,…,0) and **1**=(1,…,1). For systems with a monotonic structure the reliability function can be always computed. However, for large and complex systems this can be a hard computational task.

In order to compute the probability that the system is in the failure-free state we need to know the estimates of the elements of the vector **p**. These estimates can be obtained from the results of reliability tests. We assume that for each of system's elements we have the results of *independent* reliability tests. From these tests we obtain the vector of estimates $\mathbf{p}^* = (p_1^*, p_2^*, \ldots, p_m^*)$. The estimators p_i^* are unbiased estimators of unknown probabilities p_i only in certain particular cases. However, in the majority of practical cases, when we apply the maximum likelihood method of estimation, these estimators are asymptotically unbiased, but in practice the conditions of asymptotics usually do not hold due to the limited number of the pertinent statistical data. The knowledge of estimates $\mathbf{p}^* = (p_1^*, p_2^*, \ldots, p_m^*)$ allows for simple estimation of the reliability $R(\mathbf{p})$. In such a case we apply the method of substitution. We substitute in (4) unknown probabilities **p** with their estimates $\mathbf{p}^*$. The estimator of the reliability of the whole system $R(\mathbf{p}^*)$ is unbiased only in a particular case of systems with a series reliability structure and unbiased estimators of p_i. In all other cases $R(\mathbf{p}^*)$ is biased or at best asymptotically unbiased. Therefore, in practical situations the estimates of the system's reliability are very uncertain, and we need methods for the computation of lower bounds for its possible value. Such bounds may be obtained by the calculation of confidence intervals for $\boldsymbol{R}(\boldsymbol{p})$.

Let us now consider a system consisting of elements of *m* different types. Suppose that the reliability of the element of the *i*-th type, *i*=1,…,*m*, is a certain function of a parameter θ_i whose value is unknown. Thus, we may assume that the reliability of the whole system is described by a function $R(\boldsymbol{\theta})$ which depends on the vector $\theta = (\theta_1, \theta_2, \ldots, \theta_m)$ of parameters describing the reliability of system's elements. Moreover, we assume that the information from reliability tests of system's elements is denoted by x_i, *i*=1,…,*m*. Thus, the results of the tests are described by a vector $\mathbf{x} = (x_1, x_2, \ldots, x_m)$. We have to note that the values of θ_i and x_i only in special cases are represented by single numbers. In a general case they are represented by vectors of numbers. The interval $(\underline{R}, \overline{R})$, where $\underline{R} = \underline{R}(\mathbf{x})$ and $\overline{R} = \overline{R}(\mathbf{x})$ is the two-sided confidence interval for the unknown value of $R(\boldsymbol{\theta})$, calculated on the confidence level γ, if the following condition is fulfilled

$$P_{\boldsymbol{\theta}}\left(\underline{R} \leq R(\boldsymbol{\theta}) \leq \overline{R}\right) \geq \gamma. \quad (11)$$

In an analogical way we can define one-sided lower and upper confidence intervals for the reliability function $R(\boldsymbol{\theta})$. In the sections which follow we present methods for the calculation of such confidence intervals. In this presentation we use notation given in the book by Gnedenko *et al.* [9].

3. Confidence intervals for system's reliability in the case of discrete reliability data

Let us consider the problem of reliability estimation when the results of reliability tests of system's elements are available in a discrete form. Let us assume that the elements of all types are independently tested in exactly the same conditions as the work conditions of the considered system. In the simplest case we test samples of size N_i, *i*=1,…,*m*, for all *m* types of elements. The duration of all tests is the same, and is equal to *t*. In this simplest case we assume that we know the reliability state of each tested element at the end of the test. Thus, we assume that we know the numbers of elements d_i, *i*=1,…,*m*, which have failed during the test. The test result is described, therefore, by pairs of integer numbers (d_i,N_i), *i*=1,…,*m*. In such a case we say that these reliability tests, also known as pass-fail tests, are performed according to a *binomial scheme*. In this simple case there exists an unbiased estimator of the reliability of a tested element given by a simple formula

$$\hat{p}_i = 1 - \frac{d_i}{N_i}, i = 1, \ldots, m \tag{12}$$

The random number of the observed failures is thus described by the binomial distribution

$$P\left(d_i = d_i^*\right) = \binom{N_i}{d_i^*} \left(1 - p_i\right)^{d_i^*} p_i^{N_i - d_i^*}, i = 1, \ldots, m \tag{13}$$

Calculation of a confidence interval for the reliability p_i is not simple. For a given confidence level γone can calculate the confidence interval the formulae known as the Clopper-Pearson formulae. In the considered case of reliability estimation they have the form given in Gnedenko *et al.* [9]. The lower bound $\underline{p}$ of the one-sided confidence interval for the reliability p is given as the solution of the following equation

$$\sum_{k=0}^{d} \binom{N}{k} \left(1 - \underline{p}\right)^k \left(\underline{p}\right)^{N-k} = 1 - \gamma. \tag{14}$$

The upper bound $\overline{p}$ of the one-sided confidence interval for the reliability p is given as the solution of the equation

$$\sum_{k=0}^{N-d} \binom{N}{k} \overline{p}^k \left(1 - \overline{p}\right)^{N-k} = 1 - \gamma \quad . \tag{15}$$

In case of d=N we have $\overline{p} = 1$, and when d=0 we have $\underline{p} = 0$. It is worth noticing that if we replace $1 - \gamma$ in (14)-(15) with $0{,}5 < \alpha < 1$ and $0{,}5 < \beta < 1$, respectively, we can use these formulae for the calculation of a two-sided confidence interval for the reliability p on the confidence level equal to $1 - \alpha - \beta$.

When the probability of a failure is low i.e. when the strong inequality $q_i = 1 - p_i << 1, i = 1, \ldots, m$ holds, and when the number of tested elements N_i, i=1,…,m is large, the probability distribution of the number of failed elements d_i, i=1,…,m can be approximated by the Poisson distribution with the parameter $\Lambda_i = q_i N_i$, and the probability mass function given by the formula

$$P\left(d_i = d_i^*\right) = \frac{\Lambda_i^{d_i^*}}{d_i^*!} e^{-\Lambda_i}, i = 1, \ldots, m \tag{16}$$

This approximation is valid if for $q \to 0$ and $N \to \infty$ the condition Nq=*const* holds. One-sided confidence intervals for the parameter Λ of the Poisson distribution can be found by solving the following equations:

$$e^{-\underline{\Lambda}} \sum_{j=0}^{d-1} \frac{\Lambda^j}{j!} = 1 - \beta, \tag{17}$$

$$e^{-\overline{\Lambda}} \sum_{j=0}^{d} \frac{\Lambda^j}{j!} = \alpha \tag{18}$$

When d = 0 we have $\underline{\Lambda} = 0$. For further calculation we can use the connection between the Poisson distribution and the chi-square distribution. The confidence intervals can be thus calculated from the formulae:

$$\underline{\Lambda} = \frac{1}{2} \chi_{\beta}^2 \left(2d\right), \tag{19}$$

$$\overline{\Lambda} = \frac{1}{2} \chi_{1-\alpha}^2 \left(2d + 2\right), \tag{20}$$

where $\chi_{\gamma}^2\left(n\right)$ is the quantile of order γ of the chi-square distribution with n degrees of freedom. Similarly, as in the case of the binomial distribution, for $0{,}5 < \alpha < 1$ and $0{,}5 < \beta < 1$ we can use (19) – (20) for the calculation of the two-sided confidence interval for the parameter Λ on the confidence level $1 - \alpha - \beta$.

The Poisson distribution can be also used when the times to failure are described by the exponential distribution. When all elements failed during the test are replaced by new ones, and the duration of the test is equal to T, the observed number of failures is described by the Poisson distribution with the parameter $\Lambda = \lambda NT$, where λ is the failure (hazard) rate in the exponential distribution, and N is the number of simultaneously tested elements. Confidence intervals for the parameter Λ (and for the failure rate λ) are in this case calculated from the formulae (19) – (20).

4. Confidence intervals in the absence of observed failures

Contemporary technical systems are built of very reliable elements. For such elements we usually do not observe failures during reliability tests. In such a case, the point estimate of system's reliability is trivial, and is equal to 1. However, we are interested in the lower bound for this characteristic which may

be interpreted as kind of guaranteed reliability. Suppose, that for each of the *m* types of elements the system is built of we test N_i, i=1,…,m, elements, and in every case the number of observed failures is d_i = 0, i=1,…,m. For such test results the upper bound for the confidence interval is always equal to $\overline{R}=1$. On the other hand, it is possible to calculate the lower bound $\underline{R}$ of the confidence interval for the reliability of the considered system. In the book by Gnedenko *et al.* [9], where results of many previous works were summarized, it has been shown that the computation of this bound is equivalent to solving the following optimization problem

$$\underline{R}=\min_{\mathbf{p}\in H_0} R(\mathbf{p}), \tag{21}$$

where the set H_0 contains all values of the vector $\mathbf{p}=(p_1,p_2,\ldots,p_m)$ such that

$$\prod_{i=1}^{m} p_i^{N_i} \geq 1-\gamma \tag{22}$$

and

$$0 \leq p_i \leq 1,\ i=1,\ldots,m. \tag{23}$$

In many interesting cases there exists a closed solution to this optimization problem. In case of a series system such solution was given by Mirnyi and Solovev [19]. They showed that the lower bound of the confidence interval for system's reliability is given by a simple formula

$$\underline{R}=\min_i \underline{p}_i \tag{24}$$

where $\underline{p}_i$ is the lower bound of the one-sided confidence interval, calculated according to the Clopper-Pearson formula (14). It is easy to show that this bound can be also calculated from an equivalent formula

$$\underline{R}=(1-\gamma)^{1/N^*}, \tag{25}$$

where

$$N^*=\min_i N_i. \tag{26}$$

For systems with a more complicated structure very strong theoretical results were obtained by Pavlov [24] who considered systems with a convex cumulative risk function $H(t)$ that

$$R(t)=e^{-H(t)}. \tag{27}$$

He has shown that for such systems

$$\underline{R}=\min_i R(1,\ldots,1,\underline{p}_i,1,\ldots,1), \tag{28}$$

where

$$\underline{p}_i=(1-\gamma)^{1/N_i},\ i=1,\ldots,m. \tag{29}$$

The solutions of this problem for parallel, series-parallel, parallel-series, and *k*-out-of-*n* systems have been presented in the book by Gnedenko *et al.* [9]. For example, in the case of a system with a parallel reliability structure, consisting of *n* different elements, the lower bound of the one-sided confidence interval for system's reliability is given by:

$$\underline{R}=1-\prod_{j=1}^{n}\frac{t}{t+N_j}, \tag{30}$$

where t is the solution of the following equation:

$$\sum_{j=1}^{n} N_j \ln\left(1+\frac{t}{N_j}\right)=-\ln(1-\gamma) \quad . \tag{31}$$

In a particular case, when $N_1=\cdots=N_n=N$ Tyoskin and Kurskiy obtained a simple analytic solution (see Gnedenko *et al.* [9]) for this problem:

$$\underline{R}=1-\left[1-(1-\gamma)^{1/nN}\right]^n. \tag{32}$$

For systems with a more general coherent structure such simple solutions do not exist. However, in the book by Gnedenko *et al.* [9] two boundaries for the lower bound of the confidence interval have been proposed. Consider the set of all minimal cuts of the system, and assume that the minimal cut with the smallest number of elements consists of *b* elements. Then, consider the set of all possible minimal paths. For this set consider its all possible subsets consisting of independent, i.e. having no common elements, paths. Let *a* be the number of such paths in the subset with the largest number of independent paths. Assume additionally, that for each type of system elements exactly *N* elements have been tested. The boundaries for the lower bound for the system's reliability are the given by

$$1-\left[1-(1-\gamma)^{1/Na}\right]^a \leq \underline{R} \leq 1-\left[1-(1-\gamma)^{1/Nb}\right]^b \tag{33}$$

In a particular case of $a = b$ we have

$$\underline{R} = 1 - \left[1 - (1-\gamma)^{1/Nb}\right]^b \tag{34}$$

The authors of Gnedenko *et al.* [9] notice, that this case is typical for many reliability structures such as lattice or radial structures which are typical for large network systems.

Another very interesting method for the calculation of the lower bound of the confidence interval for system's reliability was presented in Gnedenko *et al.* [9]. Let us assume that the same vector of reliabilities $\mathbf{p} = (p_1, p_2, \ldots, p_m)$ is used for the calculation of reliability of two systems: the reliability $R(\mathbf{p})$ of the considered complex system, and the reliability $R'(\mathbf{p})$ of a simple (e.g. series) auxillary system. For this auxillary system we must know the lower bound of the respective confidence interval $\underline{R}'(\mathbf{p})$. In order to find the lower bound of the confidence interval for the reliability of the considered system we have to solve the following optimization problem:

$$\underline{R} = \min_{\mathbf{p}} R(\mathbf{p}) \tag{35}$$

where the m elements of the vector $\mathbf{p}$ must fulfill the following constraints

$$\prod_{i=1}^{m} p_i \geq \underline{R}', \; 0 \leq p_i \leq 1, \; i = 1, \ldots, m. \tag{36}$$

The lower bound calculated in this way fulfills all the requirements for a lower bound of a confidence interval, but the length of such interval is usually not the shortest possible.

5. Confidence intervals in the presence of observed failures

When failures are observed during reliability tests of system's elements the problem of building confidence intervals for the reliability of the whole system becomes much more complicated. Comprehensive information about available methods can be found in the fundamental book by Gnedenko *et al.* [9]. Below, we present only some basic results considered in this book and related literature.

Let us assume that the considered system consists of elements of m different types. For each of these types we test a sample of N_i elements, and for each sample we observe $d_i \geq 0, 1 = 1, \ldots, m$ failures. Let

$$S = R(\hat{p}_1, \hat{p}_2, \ldots, \hat{p}_m) \tag{37}$$

be the point estimator of system's reliability, where $\hat{p}_i, i = 1, \ldots, m$ are the estimators of the reliability of systems elements calculated according to (12). Now, denote by $\mathbf{d}^* = (d_1^*, d_2^*, \ldots, d_m^*)$ the vector of numbers of observed failures. Moreover, denote by $S^* = S(\mathbf{d}^*)$ the observed value of the estimator of system's reliability presented as the function of the vector $\mathbf{d}^*$. The lower bound of the confidence interval for the system's reliability is now calculated from the formula

$$\max_{\mathbf{p} \in A_R} \sum_{S(\mathbf{d}) \leq S(\mathbf{d}^*)} \prod_{i=1}^{m} \binom{N_i}{d_i} p_i^{N_i - d_i} (1 - p_i)^{d_i} = 1 - \gamma, \tag{38}$$

where maximum is calculated over the set A_R of vectors $(p_1, p_2, \ldots, p_m)$, such that

$$R(p_1, p_2, \ldots, p_m) = R, \; 0 \leq p_i \leq 1, \; 1 = 1, \ldots, m. \tag{39}$$

The sum in (38) is calculated over all possible values of the vector $\mathbf{d} = (d_1, d_2, \ldots, d_m)$ that fulfill the condition given for this sum in (38). In certain cases other formulation of this optimization problem is more suitable for computations. According to this formulation we denote by $n(\mathbf{d}) = n(d_1, d_2, \ldots, d_m)$ a non-decreasing, with respect to all components, series of vectors. The first element of this series is the vector (0,0,…,0), and then we have the vectors of the type (0,..,0,1,0,…,0), etc. The lower bound of the confidence interval for system's reliability can be calculated from

$$\underline{R} = \min R(p_1, p_2, \ldots, p_m), \tag{40}$$

where minimum is taken over the set of all values of the vector $(p_1, p_2, \ldots, p_m)$ such that

$$\sum_{n(\mathbf{d}) \leq n(\mathbf{d}^*)} \prod_{i=1}^{m} \binom{N_i}{d_i} p_i^{N_i - d_i} (1 - p_i)^{d_i} \geq 1 - \gamma, \quad 0 \leq p_i \leq 1, \; i = 1, \ldots, m \tag{41}$$

The optimization problem given by (40) – (41) was formulated first time by Buehler [6] who considered a system consisted of two elements. This was the first result of the calculation of the confidence interval for system's reliability.

Let us now consider the series system consisted of m different elements. The optimization problem is now the following:

$$\underline{R} = \min \prod_{i=1}^{m} p_i \,, \tag{42}$$

where minimum is taken over all vectors $(p_1, p_2, \ldots, p_m)$ such that

$$\sum_{R(\mathbf{d}) \geq R\left(\mathbf{d}^*\right)} \prod_{i=1}^{m} \binom{N_i}{d_i} p_i^{N_i - d_i} (1 - p_i)^{d_i} \geq 1 - \gamma, \tag{43}$$

$$0 \leq p_i \leq 1, i = 1, ..., m$$

The calculation of the lower bound of the confidence interval for system's reliability $\underline{R}$ can be simplified when the probabilities of failures are small, i.e. when the inequality $q_i = 1 - p_i << 1, i = 1, \ldots, m$ holds. In such a case we can assume that the number of failures is described by the Poisson distribution with the parameter $\Lambda_i = q_1 N_i, i = 1, \ldots, m$. It has been shown in the book by Gnedenko *et al.* [9] that in this case we have

$$\underline{R} = e^{-\bar{f}} \tag{44}$$

where

$$\bar{f} = \max\left(\sum_{i=1}^{m} \frac{\Lambda_i}{N_i} \right), \tag{45}$$

and the maximum in (45) is taken over all vectors $\Lambda = (\Lambda_1, \Lambda_2, \ldots . \Lambda_m)$ such that

$$\sum_{R(\mathbf{d}) \geq R\left(\mathbf{d}^*\right)} \prod_{i=1}^{m} e^{-\Lambda_i} \left(\frac{\Lambda_i^{d_i}}{d_i !} \right) \geq 1 - \gamma, \lambda_i \geq 0, \tag{46}$$

$$i = 1, ..., m$$

This practical result was obtained first time by Bol'shev and Loginov [5] for the case of equal values of N_i, and, independently, by Pavlow [23] and Sudakov [25] for any values of these numbers.

6. Approximate confidence intervals for system's reliability

Computation of exact bounds of confidence intervals for system's reliability requires, with only few exceptions, solving difficult optimization problems. Therefore, its practical applicability is somewhat limited unless specialized software is available. For this reason several authors, mainly American, have tried to obtain approximate, but relatively easy for computation, solutions. Different approximate solutions have been proposed by such authors as Madansky [12], Myhre and Saunders [20], Easterling [7], Mann [13],[14], or Mann and Grubbs [15],[16]. Comprehensive review of such results can be found in a well known book by Mann, Shaefer, and Singpurwalla [17]. However, probably the most interesting from a practical point of view result was presented in one of the first textbooks on reliability written by Lloyd and Lipow [11]. These authors presented a heuristic method, attributed to Lindstrom and Madden, for the calculation of the approximate confidence interval for the system with a series reliability structure. This method utilizes the concept of so called equivalent tests. To present this method we consider, following the book by Gnedenko *et al.* [9], a system with a series-parallel structure which has the same elements in its parallel subsystems. Let R^* be the estimated value of the reliability function for the considered system, and N_i, i=1,…,m be the number of tested items for the element of the i-th type. The equivalent number of failures D_i^* for the element of this type is then calculated from the equation

$$R\left(1, \ldots, 1, 1 - D_i^* \Big/ N_i, 1, \ldots, 1\right) = R^* \tag{47}$$

At the next stage of the computation procedure, for each equivalent test (N_i, D_i^*) we calculate the lower bound of the confidence interval $\underline{P}_i(N_i, D_i^*)$ by solving the equation

$$B_p\left(N_i - D_i^*, D_i^* + 1\right) = 1 - \gamma, \tag{48}$$

where

$$B_p(a,b) = \frac{\int_0^p x^{a-1}(1-x)^{b-1} dx}{\int_0^1 x^{a-1}(1-x)^{b-1} dx} \tag{49}$$

is the incomplete beta function whose values can be computed using available numerical procedures. The lower bound of the confidence interval is now calculated from a simple formula

$$\underline{R} = \min_{1 \leq i \leq m} R\left(1, \ldots, 1, \underline{P}_i\left(N_i, D_i^*\right), 1, \ldots, 1\right). \tag{50}$$

The Lindstrom-Madden method was proposed as an approximate heuristic method. However, it has been proved (see the book by Gnedenko *et al.* [9] for additional information) that for many simple

reliability structures it produces exact confidence intervals.

Another method which uses the concept of equivalent tests, and which can be used for the analysis of complex systems consisted of many simple subsystems, was proposed by Martz and Duran [18]. In this method it is assumed that for each simple subsystem we are able to calculate the value of its reliability estimator R_i, and the lower bound for the respective confidence interval $\underline{R}_i$. Next, from a set of equations

$$1-\frac{r_i}{M_i}=R_i \tag{51}$$

and

$$\underline{R}_i=\underline{P}_i\left(M_i,r_i\right) \tag{52}$$

we calculate the parameters (M_i,r_i) of the equivalent binomial reliability tests. In further analysis the considered subsystem is treated as a single element described by the equivalent test. Note, that for the application of this method it is not important how we have found the values of R_i and $\underline{R}_i$.

7. Some remarks about other methods for the calculation of confidence intervals for system's reliability

In the previous sections we have presented methods for the calculation of confidence intervals for system's reliability for the case of discrete reliability data from tests, i.e. when the numbers of tested elements and the numbers of observed failures are known. It is a well known fact that the knowledge of lifetime distributions combined with the knowledge of observed times to failures may increase the accuracy of reliability estimation. Moreover, this knowledge may be sufficient for the prediction of reliability at time instants other than the times of the performed reliability tests. Unfortunately, even in the simplest case of the exponential distribution of lifetimes the exact and practically applicable solutions are known only in few cases when lifetime tests are performed according to the type-II censoring scheme (a fixed number of observed failures). For example, Lentner and Buehler [10] considered the case of a series system with only two elements. Their result was generalized in an unpublished PhD thesis by El Mawaziny [8] who proposed an iterative method for the calculation of the lower bound of the confidence interval for reliability of a series system consisted of m elements. Because of its complicated nature this algorithm has not been described in reliability textbooks. However, there exists a good approximation proposed by Mann and Grubbs [15], and in a simplified version by Mann [14].

Consider the case when the lifetimes are exponentially distributed, and reliability tests provide type-II censored data. For each type of system elements we test a sample of n_i items, and observe times $t_{i,j}$ of the first $r_i>0$, $i=1,\ldots,m$ failures. The respective value of the total time on test z_i, is given by

$$z_i=\sum_{j=1}^{r_i}t_{i,j}+\left(n_i-r_i\right)t_{i,r_i},\ i=1,\ldots,m \tag{53}$$

Denote by $z_{(1)}$ the minimal value of z_i, $i=1,\ldots,m$. Mann [14] has shown that the estimator of the hazard rate of the series system has approximately the expected value given by

$$\mu=\sum_{i=1}^{k}\frac{r_i-1}{z_i}+\frac{1}{z_{(1)}}, \tag{54}$$

and the variance given by

$$\nu=\sum_{i=1}^{k}\frac{r_i-1}{z_i^2}+\frac{1}{z_{(1)}^2}. \tag{55}$$

To approximate the optimum lower bound for reliability of a series system $\underline{R}_s(t)$ at confidence level β, using the Wilson-Hilferty transformation, one calculates

$$\underline{R}_s\left(t\right)=\exp\left[-t\mu\left(1-\frac{\nu}{9\mu^2}+\frac{y_\gamma\nu^{1/2}}{3\mu}\right)^3\right], \tag{56}$$

where y_γ is the quanitile from the standardized normal distribution. For systems with more complex structures an interesting approach has been proposed in Gnedenko *et al.* [9]. According to this approach first we have to calculate upper bounds for the hazard rates of system's elements using the following simple formula

$$\overline{\lambda}_i=\frac{\chi_\gamma^2\left(2r_i\right)}{2S_i},\ i=1,\ldots,m, \tag{57}$$

where $\chi_\gamma^2\left(2r\right)$ is the quantile of the γ order from the chi-square distribution with $2r$ degrees of freedom. When we insert these lower bounds into a formula

for the calculation of the system's reliability function instead of respective hazard rates, i.e. if we calculate

$$\underline{R} = R\left(\overline{\lambda}_1.\overline{\lambda}_2,\ldots,\overline{\lambda}_m\right), \tag{58}$$

the obtained value usually fulfills the requirements for a confidence interval. Pavlov [24] has shown that in case of γ>0,778 this approach allows to calculate confidence intervals for a broad class of reliability structures for lifetime distributions having non-decreasing (in time) hazard rates (i.e. for elements with the ageing property).

The general methodology for the calculation of confidence intervals for system's reliability was proposed by Belyaev [3], [4]. Other, but completely equivalent general method, was proposed by Bol'shev and Loginov [5]. Below, we present the main results of Belyaev.

Suppose that we know the statistic S which can be used as a point estimator of system's reliability, i.e. $S = \hat{R}$. Moreover, we assume that this statistic is a function of a vector of parameters $\boldsymbol{\theta}$ describing probability distributions of lifetimes of system's elements. Additionally, we assume that the probability distribution of this statistic *is known*, i.e. we know

$$F(t,\boldsymbol{\theta}) = P_{\boldsymbol{\theta}}(S \le t). \tag{59}$$

For a given value of the vector $\boldsymbol{\theta}$ we can now introduce two functions $t_1(\boldsymbol{\theta})$ and $t_2(\boldsymbol{\theta})$, such that

$$F(t_1,\boldsymbol{\theta}) = \alpha \tag{60}$$

and

$$F(t_2,\boldsymbol{\theta}) = 1-\beta. \tag{61}$$

Now, let's denote by

$$A_R = \{\boldsymbol{\theta} : R(\boldsymbol{\theta}) = R\} \tag{62}$$

the set of all values of the vector $\boldsymbol{\theta}$ for which the reliability function adopts a given value R. Next, introduce two functions

$$K_1(R) = \min_{\boldsymbol{\theta}\in A_R} t_1(\boldsymbol{\theta}) \tag{63}$$

and

$$K_2(R) = \min_{\boldsymbol{\theta}\in A_R} t_2(\boldsymbol{\theta}). \tag{64}$$

The lower and upper bounds of the confidence interval on the confidence level $1-\alpha-\beta$ for system's reliability can be found by solving equations

$$K_1(\underline{R}) = S^* \tag{65}$$

and

$$K_2(\overline{R}) = S^*, \tag{66}$$

where S^* is the observed value of the statistic S.

The described general methodology is based on the original methodology for the construction of confidence sets proposed by Neyman [22], and is valid for any type of reliability data, and any reliability structure. However, its practical applicability is limited only to rather simple cases.

8. Approximate lower bounds for system's reliability based on minimum values of the reliability of system's elements

Computation of optimal (i.e. the shortest) and exact confidence intervals is, with a few exceptions, a very difficult task. Moreover, in all published results it is assumed that the elements in a system are mutually independent. Additional problems arise from a fact that confidence intervals used for the description of test results may be conservative, as in the case of intervals based on the Clopper-Pearson formula. In this section we present approximate bounds for system's reliability which, under certain conditions, may replace lower bounds of confidence intervals.

In order to investigate the robustness of the confidence intervals for system's reliability against the departure from the assumption of independence of system's elements let us introduce the notion of a *copula*. According to a famous theorem of Sklar (see e.g. Nelsen [21]) any two-dimensional probability distribution function $H(x,y)$ with marginals $F(x)$ and $G(y)$ is represented using a function C, called a *copula*, in the following way:

$$H(x,y) = C(F(x),G(y)) \tag{67}$$

for all $x, y \in R$. Conversely, for any distribution functions F and G and any copula C, the function H defined by (67) is a two-dimensional distribution function with marginals F and G. Moreover, if F and G are continuous, then the copula C is unique. In our investigation we have considered three types of copulas:

a) Clayton copula, defined as

$$H(x,y) = \left[F^{-\theta}(x) + G^{-\theta} - 1\right]^{-1/\theta}, \theta > 0 \qquad (68)$$

b) Gumbel copula, defined as

$$H(x,y) = \exp\left(-\left[(-\ln F(x))^{\theta} + (-\ln G(y))^{\theta}\right]^{1/\theta}\right), \qquad (69)$$
$$\theta > 0$$

c) Fairlie-Gumbel-Morgenstern (FGM) copula defined as

$$H(x,y) = F(x)G(y)(1+\theta(1-F(x))(1-G(y))), \qquad (70)$$
$$-1 \le \theta \le 1$$

The Clayton and Gumbel copulas can be used for modeling a positive stochastic dependence. The FGM copula can be used for modeling both negative (θ<1) and positive (θ>1) dependence. The Clayton copula is especially interesting in reliability applications as it describes stronger dependence for smaller lifetimes than for larger ones. If this type of dependence exists the reliability of a series system with dependent elements is greater than in the case of independence. On the other hand, for a parallel system the reliability of a system with dependent elements is smaller.

In the majority practical cases the reliability of tested elements is high, and even for moderate sample sizes the number of observed failures is small. This suggests utilization of the result obtained for the case of zero-failure tests for the calculation of the lower bounds for reliability of a series system given by the expression (24). To analyze the properties of this approximation let us consider a two-element series system whose elements are equally reliable. We also assume that the sample sizes for both elements are the same. On Figure 1 we present the comparison of the values of our simple approximate bound with the bounds calculated for this system using a substitution method. For obtaining the presented results we performed a Monte Carlo simulation experiments, and in each of them we generated 500 000 test cases, Our approximate bound, plotted against the expected number of observed failures in a sample (for a probability of failure equal to 0,01), is represented by a continuous upper curve. The middle curve represents the bound calculated by the insertion into (5) the respective lower bound of the confidence intervals for the reliability of elements, calculated for the same confidence level (γ=0,9).The lower curve is a similar to the previous one, but calculated for the confidence level equal to $\sqrt{\gamma}$, as it is suggested in statistical literature.

For these bounds we have estimated the coverage probability of the considered confidence intervals. The results of the comparison are presented on *Figure 2* for our approximate bound, and the bound represented by the middle curve on *Figure 1*.

As we can see, our simple bound fulfills requirements for a confidence interval not only for zero-failure reliability tests, but for all tests with the expected number of failures not greater than 1,95. The classical and much wider confidence intervals have the probability of coverage close to 1, i.e. much greater than the designed value of 0,9.

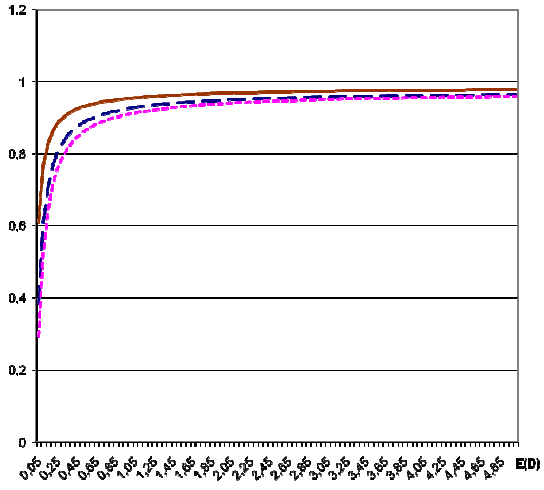

Figure 1. Lower bounds for a series system

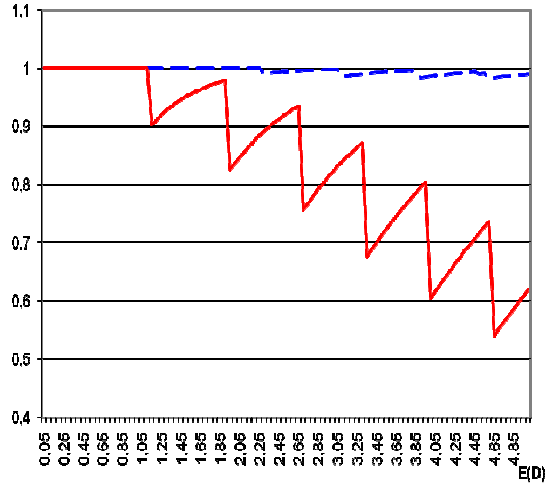

Figure 2. Coverage probabilities for a series system in case of independence

Now, let us consider the case when the elements of the system are dependent. On *Figure 3* we show the coverage probability when this dependence is described by the Clayton copula with dependence

parameter θ=2, and the Gumbel copula, with dependence parameter θ=2. For this value of the parameter the Kendall measure of dependence τ for both copulas is equal to 0,5. It means that the dependence is positive and fairly strong. The coverage probability in the case of the Clayton copula (solid line) is greater than the designed value for tests with the expected value of observed failures greater than 5. However, in the case of the dependence described by the Gumbel copula (dashed line) this feature is guaranteed only for this value not greater than 2. It shows, how the type of dependence influences the results despite the fact that the popular measure of dependence, such as Kendall τ in both cases gives exactly the same value. Similar results have been also obtained for the FGM copula which can be used for modeling weaker dependencies.

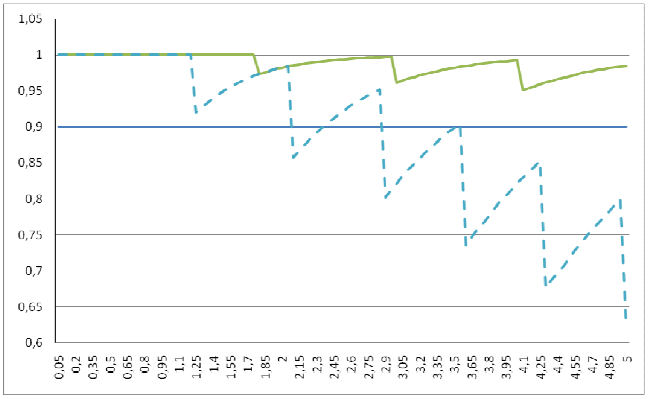

Figure 3. Coverage probabilities for a series system in case of dependence

Now, let us consider the case of the system with elements connected in parallel. For such systems a simple for computation bound which is similar to that for a series system does not exist. Instead we propose the following approximation

$$\underline{R} = 1 - \min_i \overline{q}_i \quad , \tag{71}$$

where $\overline{q}_i$ is the upper bound of the confidence interval for the probability of failure. The lower bound calculated according to (71) is *always* smaller than the bound obtained by substitution of the probabilities of failures q_i with their respective upper bounds $\overline{q}_i$. Thus, the coverage probability in case of independent elements of the system, calculated according to (71), is always greater than the respective confidence level. It can be seen at *Figure 4*, where this probability is always equal to 1. (Note that the coverage probability in case of the bound obtained by substitution is also much greater than the confidence level which is equal to 0,9).

The situation changes dramatically when the elements of the system are positively dependent, e.g. when their dependence is described either by the Clayton copula or by the Gumbel copula. On *Figure 5* we present the estimated coverage probabilities in such cases when the confidence intervals are calculated using the substitution method.

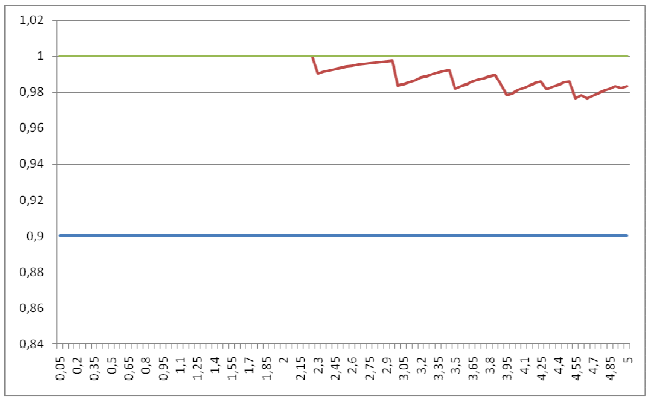

Figure 4. Coverage probabilities for a parallel system in case of independence

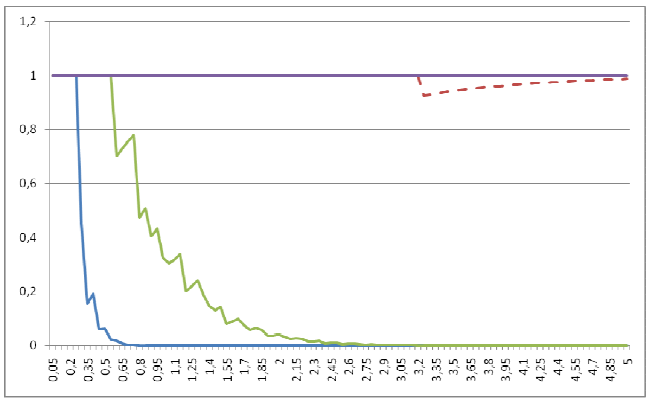

Figure 5. Coverage probabilities for a parallel system in case of dependence

The coverage probabilities (the left-most curve for the Clayton copula, and the curve next to it for the Gumbel copula) show dramatically that the confidence intervals obtained by substitution under the assumption of independence are too narrow. On the other hand, the interval calculated according to (71) has the coverage probability (depicted by a dashed curve for the Clayton copula, and equal to one for the Gumbel copula) greater than the confidence level.

9. Conclusion

Many prominent authors, mainly from USA and the Soviet Union, contributed to the problem of computing the lower confidence bounds for system's reliability using the data from tests of separate elements or subsystems. The proposed exact bounds

are usually difficult to compute. Good approximations exist, but they are usually obtained under the assumption that failures of all elements or subsystems are observed during the tests. In the paper we have shown using Monte Carlo simulation that in case when elements working together in a system are dependent these bounds are inaccurate or even useless, as it is the case of parallel (redundant) systems. In the paper, we have proposed very simple bounds characterized by satisfactory performance, at least for highly reliable system elements, which are robust against the presence of positive dependence of the elements of a system. The results of simulation experiments indicate that the proposed approach is also applicable for series-parallel and parallel-series systems defined by (7) and (8), respectively. However, the experiments also show that our approximations have worse properties for systems with larger number of elements, unless these elements are very reliable.

References

[1] Barlow, R. E. & Proschan, F. (1965). *Mathematical Theory of Reliability*. John Wiley, New York.

[2] Barlow, R. E. & Proschan, F. (1975). *Statistical Theory of Reliability and Life Testing. Probability Models*. Holt Rinehart and Winston, Inc., New York.

[3] Belyaev, Y. K. (1966). Confidence intervals for the functions of many unknown parameters. *Proc. of the Sov. Acad. Sci*, 169, 755-758 (in Russian)

[4] Belyaev, Y. K. (1968). On simple methods of Confidence Limits Construction. *Eng. Cybernetics*, No.5, 96-101 (in Russian, exists English translation of this journal)

[5] Bol'shev, L.N. & Loginov E.A. (1966). Interval estimates under noisy parameters. *Theory of Prob. and its Appl.* v.11, No.1, 94-107 (in Russian, exists English translation of this journal)

[6] Buehler R.J. (1957). Confidence Intervals for the Product of two Binomial Parameters. *Journal of the American Statistical Association*, 52, 482-493.

[7] Easterling, R.G. (1972). Approximate Confidence Limits for the System Reliability. *Journal of the American Statistical Association*, 67, 220-222.

[8] El Mawaziny, A.H. (1965), *Chi-square distribution theory with applications to reliability problems*. PhD Thesis, Iowa State Univ., Ames, Iowa.

[9] Gnedenko, B., Pavlov, I., Ushakov, I. & Chakravarty S. (1999). *Statistical Reliability Engineering*. John Wiley, New York.

[10] Lentner, M.M. & Buehler, R.J. (1963). Some inferences about gamma parameters with an application to a reliability problem. *Journal of the Amer. Stat. Assoc.* 58, 670-677.

[11] Lloyd, D.K. & Lipow, M. (1962). *Reliability Management. Methods and Mathematics*. Prentice-Hall, Englewood Cliffs, NJ.

[12] Madansky, A. (1965). Approximate Confidence Limits for the Reliability of Series and Parallel Systems. *Technometrics*, 7, 493-503.

[13] Mann, N.R. (1974). Approximate Optimum Confidence Bounds on Series and Series-Parallel System Reliability for Systems with Binomial Subsystem Data. *IEEE Transactions on Reliab*ility, R-23, 295-304.

[14] Mann, N.R. (1974). simplified expressions for obtaining approximately optimum system-reliability confidence bounds from exponential subsystem data. *Journal of the Amer. Stat. Assoc*., 69, 492-495

[15] Mann, N.R. & Grubbs, F.E. (1972). Approximately Optimum Confidence Bounds on Series-System Reliability for Exponential Time to Fail Data. *Biometrika*, 50, 191-204.

[16] Mann, N.R. & Grubbs, F.E. (1974). Approximately Optimum Confidence Bounds for System Reliability Based on Component Test Plan. *Technometrics*, 16, 335-347

[17] Mann, N.R., Schaefer, R.E. & Singpurwalla, N.D. (1974). *Methods of Statistical Analysis of Reliability and Life Data*. John Wiley, New York.

[18] Martz, H.F. & Duran, I.S. (1985). A Comparison of Three Methods for Calculating Lower Confidence Limits on System Reliability Using Binomial Component Data. *IEEE Transactions on Reliability*, R-34, 113-120.

[19] Mirnyi, R.A. & Solovyev, A.D. (1964). Estimation of system reliability on the basis on its units tests. In: *Cybernetics in Service for Communism*, vol.2, Energiya, Moscow, 213-218.

[20] Myhre, J. & Saunders, S.C. (1968). Comparison of Two Methods of Obtaining Approximate Confidence Intervals for System Reliability. *Technometrics*, 10, 37-49.

[21] Nelsen, R.B. (2006). *An Introduction to Copulas* (2nd edition), Springer, New York.

[22] Neyman, J. (1935). On the Problem of Confidence Intervals. *Annals of Mathematical Statistics*, 6, 111-116.

[23] Pavlov, I.V. (1973). Confidence limits for system reliability on the basis on its components testing. *Eng. Cybernetics*, No.1, 52-61 (in Russian, exists English translation of this journal)

[24] Pavlov, I.V. (1982). *Statistical methods of reliability estimation by tests results*. Radio i Svyaz, Moscow (in Russian).
[25] Sudakov, R.S. (1974). About interval estimation of reliability of series system. *Eng. Cybernetics*, No.3, 86-94 (in Russian, exists English translation of this journal).