

Chumakov I.A., Chepurko V.A., Antonov A.V.

ON SOME PROPERTIES OF KIJIMA INCOMPLETE RECOVERY MODELS

The article analyses some properties of Kijima incomplete recovery models using a Weibull distribution for time to first failure. The maximum likelihood method is used for assessment of distribution parameters and recovery coefficient. Confidence limits have been identified using a Fisher information matrix. The authors consider cases of processing data from several identical elements and prove the inverse relationship between the deviation value and the number of elements. The paper examines two ways of assessing the leading function of the deteriorating component flow. A comparison is made between the new approach that represents the leading function of the flow as the ultimate sum and the approach that uses the statistical testing method. The paper suggests the method of calculation of the average direct time and reverse residual time based on the statistical testing method. Several demonstration examples are given.

Keywords: *incomplete recovery, Kijima models, Weibull distribution, leading function of flow, average direct residual time.*

Introduction

Models that involve complete or minimal recovery are the most commonly used in dependability calculation of recoverable systems. Technical systems, as a rule, normally function in a more complex manner with incomplete (partial) recovery. Models that take into consideration incomplete recovery are becoming more and more popular. That includes the Kijima models that are covered in this paper.

1. Extended recovery process, Kijima models

In case of immediate recovery the real age of an element at the moment of the n-th recovery can be represented as the sum of all of its times to failure:

$$S_n = \sum_{i=1}^{n} X_i, \qquad S_0 = 0;$$

where X_i is the time to the i-th failure.

Let us introduce a certain constant value q that is called a recovery coefficient (factor). We will define the virtual age of an element as a certain function v such that $v = v(\{X\}, q)$. The virtual age and distribution of time to failure are related by the following formula: let v_{i-1} be the virtual age of the element at the moment of the $(i-1)$-th recovery. Then, random value X_i has the following conventional distribution function [1, 2]:

$$F_i(x \mid v_{i-1}) = \frac{F(x+v_{i-1}) - F(v_{i-1})}{1 - F(v_{i-1})}; \tag{1}$$

where $F(x)$ is the function of distribution of time to first failure (TFF) for an absolutely new element.

The Kijima-1 model implies that the n-th recovery affects only the damage received by an element between the ($n-1$)-th and the n-th failures reducing the element's virtual age increment from X_i to qX_i. An element's virtual age after the n-th recovery can be written as follows:

$$v_n = v_{n-1} + qX_n = q\sum_{i=1}^{n} X_i = qS_n;\ v_0 = 0; \qquad (2)$$

The Kijima-2 model implies that each recovery affects the total damage, thus reducing the total virtual age:

$$v_n = qv_{n-1} + qX_n = q(q^{n-1}X_1 + q^{n-2}X_2 + \ldots + X_n);\ v_0 = 0; \quad (3)$$

Therefore, the TFF distribution and coefficient q completely define the Kijima models recovery processes. Among others, according to [2, 3, 4] the case of $q = 0$ describes a complete recovery, the case of $q = 1$ describes a minimal recovery, the case of $0 < q < 1$ describes incomplete recovery "worse than new but better than before the failure". Model parameters can be evaluated in a number of ways.

2. Parametric evaluation of model parameters

2.1 Method of maximum likelihood

Let us consider the approach based on the method of maximum likelihood (MML) [1, 5]. Here and further the assumption is that the TFF has a Weibull distribution. The function of this distribution has various notations. In this paper, the following form is used in order to simplify further calculations:

$$F(x) = 1 - \exp(-\lambda x^{\beta});\quad \beta \geq 1. \qquad (4)$$

For (1) subject to (4) the log-likelihood function (LLF) is known [4]:

$$\ln L(\lambda,\beta,q) = n(\ln\lambda + \ln\beta) + \lambda\sum_{i=1}^{n}[v_{i-1}{}^{\beta} - (X_i + v_{i-1})^{\beta}] +$$

$$+(\beta-1)\sum_{i=1}^{n}\ln(X_i + v_{i-1}); \qquad (5)$$

where v_i depend on q as (2) or (3).

Estimation л, в, q can be derived by means of numerical techniques [1, 5].

Then let us assume that under observation are simultaneously k absolutely identical recoverable elements. In this case we assume that the i-th failure times have identical distributions for each element. Let us write LLT as in (5), yet taking into consideration k samples of failure times [3]:

$$\ln L(\lambda,\beta,q) = \sum_{j=1}^{k} n_j(\ln\lambda + \ln\beta) +$$

$$+\lambda\sum_{j=1}^{k}\sum_{i=1}^{n}[v_{j,i-1}{}^{\beta} - (X_{j,i} + v_{j,i-1})^{\beta}] + (\beta-1)\sum_{j=1}^{k}\sum_{i=1}^{n}\ln(X_{j,i} + v_{j,i-1}). \ (6)$$

Function (6) allows finding parameter estimations as in (5).

2.2 Error estimation of the method of maximum likelihood

In order to find the MML estimation variance, we must use a Fisher information matrix [6, page 201]. Let us write the parameter vector (л, в, q) as (и$_1$, и$_2$, и$_3$). The elements of a Fisher information matrix are calculated as follows:

$$I(i,j) = -M[\frac{\partial^2 \ln L}{\partial\theta_i \partial\theta_j}].$$

Estimate valiances are calculated by covariance matrix $V = I^{-1}$, while D(и$_i$)=$V(i,i)$. For LLT (6) of k observed identical elements it is possible to similarly calculate elements $I_k(i,j)$ of the Fisher matrix I_k, then to find the respective variance D_k(и$_i$).

Any j-th of the observed k elements has the number of failures n_j with identical distribution and mathematical expectation (ME):

$$Mn_j = Mn_1;\ j = 1\ldots k.$$

For the i-th time to failure of the j-th element $X_{j,i}$ and respective virtual age of $v_{j,i-1}$ we deduce:

$$MX_{j,i} = MX_{1,i};\quad j = 1\ldots k;$$

$$Mv_{j,i-1} = Mv_{1,i-1};\quad j = 1\ldots k.$$

Also note that the times before failure of the first element $X_{1,i}$ do not depend on the times before failure of the second element $X_{2,i}$, third, etc. for any i. Let us examine the elements of the matrix I_k constructed for (6) and compare them with the elements of the matrix I_1 constructed for one element using (5):

$$I_k(1,1) = -M[\frac{\partial^2 \ln L}{\partial\lambda^2}] = -M[\frac{\sum_{j=1}^{k} n_j}{\lambda^2}] =$$

$$= -kM[\frac{n_1}{\lambda^2}] = kI_1(1,1) = kI(1,1).$$

Similarly, for all the other elements $I_k(i,j) = kI_1(i,j)$, from where we deduce the variance:

$$D_k(\theta_i) = V_k(i,i) = I_k^{-1}(i,i) = \frac{1}{k}I_1^{-1}(i,i) =$$

$$= \frac{1}{k}D_1(\theta_i) = \frac{1}{k}D(\theta_i).$$

The results indicate that in statistical terms **the estimate variance is negatively related to** the number of observed identical elements. In other words, estimate variance per k samples of the same size will be k times smaller than the estimate variance per a single sample. It should be noted that Kijima processes are not homogenous, therefore this property does not follow from the definition of the processes themselves.

3. Estimation of the leading function of the stream based on sum total calculation

The ME of the mean failures per interval (0; t] is also known as the recovery function, leading function of the flow (LFF). For the case of incomplete recovery this value is defined with an integral equation that is impossible to solve analytically, while even the computational solution is complicated [2]. By definition, LFF can be represented as an infinite sum [7, стр.88]:

$$H(x) = \sum_{i=1}^{\infty} G_i(x), \tag{7}$$

where $G_i(x)$ is the distribution function of the time point of the i-th failure of S_i.

Unlike distribution (1), $G_i(x)$ is an unconditional distribution. Let us find $G_i(x)$. In order to do that let us consider S_i as a sum of random values S_{i-1} and X_i, where X_i is the i-th time to failure. Next, let us find the distribution function of this sum:

$$\begin{aligned} G_i(x) &= \mathrm{P}\left(S_{i-1} + X_i < x\right) = \\ &= \int_0^x g_{i-1}(y) \int_0^{x-y} \frac{f(z+qy)}{1-F(qy)} dz dy = \\ &= \int_0^x \frac{g_{i-1}(y)}{1-F(qy)} [F(x+y(q-1)) - F(qy)] dy. \end{aligned} \tag{8}$$

Let us transform (8) in order to make it suitable for calculations:

$$G_i(x) = \int_0^x G_{i-1}(y) R(x,y) dy, \tag{9}$$

$$\begin{aligned} R(x,y) = &-\frac{(q-1) f(x+y(q-1))}{1-F(qy)} - \\ &-\frac{q f(qy)}{(1-F(qy))^2}[F(x+y(q-1))-1]. \end{aligned}$$

In case of a Weibull distribution the last equation changes to:

$$R_W(x,y) = \lambda\beta(q(qy)^{\beta-1} - (q-1)d^{\beta-1}) \exp(\lambda((qy)^{\beta} - d^{\beta})),$$

where $d = x + y(q - 1)$.

Equation (9) defines the recurrent dependence between the distribution functions of failure time points. Knowing the TFF distribution (4) we can calculate a certain sum total of the distribution function. A similar recurrence equation was deduced for distribution density of the i-th time to failure:

$$\begin{aligned} g_i(x) = \frac{\partial F_i(x)}{\partial x} &= \int_0^x \frac{g_{i-1}(y)}{1-F(qy)} f(x+y(q-1)) dy = \\ &= \int_0^x g_{i-1}(x) Q(x,y) dy. \end{aligned}$$

In case of a Weibull distribution the equation $Q(x,y)$ changes to:

$$Q_W(x,y) = d^{\beta-1} \exp(\lambda((qy)^{\beta} - d^{\beta})).$$

Out of (7) we can find the value of LFF with a certain error that decreases as i increases.

4. Evaluation by means of statistical tests

4.1 Modeling of times to failure

Assuming that the distribution parameters and recovery coefficients of the Kijima models known, it becomes possible to model Kijima processes (2) and (3). The TFF distribution function is found as (1). Out of [4] we will deduce the formula for the i-th time to failure:

$$X_i = ((v_{i-1})^{\beta} - \lambda^{-1} \ln U)^{\beta^{-1}} - v_{i-1}. \tag{10}$$

For modelling, $U \sim U[0;1]$ must be played.

4.2 Estimation of the leading function of the flow

In case of incomplete recovery, a popular way of calculating LFF is the method of statistical tests that, among others, is suitable for evaluation prediction in the future. In order to evaluate LFF in point t, it is required to play a sequence of random values, times to failure. N_j is the number of failures that occurred within time t, i.e. the value of LFF. Then, modeling is repeated the required number of times S. The final estimate of LFF $H_M(t)$ is calculated as the mean of the modeled values [5]:

$$H_M(t) = \frac{1}{S} \sum_{i=1}^{S} N_i. \tag{11}$$

4.3 Estimation of the average direct time and reverse residual time

The average direct residual time (ADRT) [7] is the ME of the remaining time of facility operation till the next failure from the time point t when the system was operable.

The average reverse residual time (ARRT) is the ME of facility operation time from the beginning of operation or last recovery till the time point t when the system is operable.

Those life characteristics are calculated only for recoverable elements. Similarly to LFF, a direct calculation of ARRT and ARRT for Kijima models is not a trivial task, therefore the authors propose an approach based on statistical testing. In order to evaluate ADRT and ARRT in point t, it is required to play a sequence of random values , times to failure $\{X_n\}$, that corresponds to the sequence of failure points $\{T_n\}$ such that $T_{n-1} < t \le T_n$ as shown in Figure 1.

Estimates of ADRT $V(t)$ and ARRT $R(t)$ are calculated as the average of the modeled values:

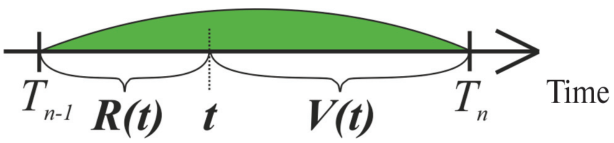


Fig. 1. Modeling of failure points for calculation of residual time

$$V(t) = \frac{1}{S}\sum_{i=1}^{S}(T_n - t); \quad (12)$$

$$R(t) = \frac{1}{S}\sum_{i=1}^{S}(t - T_{n-1}). \quad (13)$$

Item 6 below will list the results of calculations using those formulas.

4.4. Evaluation of calculation errors by means of statistical tests

As it is known, the method of statistical testing enables error evaluation that only has a certain degree of confidence. According to [8, page 234] and CLT we deduce the upper limit of error for (11) with confidence coefficient в:

$$\delta \le t_\beta \sqrt{\frac{D(H_M)}{S}}; \quad (14)$$

where $t_в$ is the value of argument of Laplace's function $\Phi(t)$ wherein $\Phi(t)$ = в/2.

$$D(H_M) = \frac{\sum_{i=1}^{S}(H_M - N_i)^2}{S-1}$$

is the unbiased estimator of variance of estimate H_M.

5. Investigation of the existence of the accumulation point of the Kijima process

5.1 Divergence of failure points sequence for minimal recovery

For some incomplete recovery models the sequence **of mathematical expectations** (ME) of time points of the i-th failure $M(S_n)$ can converge, i.e. have a limit under $i \to \infty$. In particular, for a geometrical process, a sequence of non-negative independent random values {$Д_n$; $n = 1, 2\ldots$} such that the following equation for distribution [9, page 81] is correct:

$$\Delta_{n+1} \overset{d}{=} \gamma \Delta_n;$$

the following is correct:

$$M(S_\infty) = M\left(\lim_{n\to\infty} S_n\right) = \frac{M\Delta}{1-\gamma};$$

where $MД$ is the ME of the first time to failure; г > 0 is the denominator (parameter) of the geometrical process.

Let us prove the absence of convergence of the ME of the sequence failure points for Kijima models for a Weibull distribution. Let us consider the special case $q = 1$. Formula (9) changes to:

$$G_i(x) = \int_0^x G_{i-1}(y)\lambda\beta y^{\beta-1}\exp(\lambda(y^\beta - x^\beta))dy =$$

$$= e^{-\lambda x^\beta}\int_0^x e^{\lambda y^\beta} G_{i-1}(y)d(\lambda y^\beta) =$$

$$= \left\{\lambda y^\beta = t;\ y = \left(\frac{t}{\lambda}\right)^{\frac{1}{\beta}} = U;\ \lambda x^\beta = A\right\} =$$

$$= e^{-A}\int_0^A e^t G_{i-1}(U)dt;$$

First, let us note:

$$G_1(x) = 1 - e^{-A}; G_1(U) = F_1\left[\left(\frac{t}{\lambda}\right)^{\frac{1}{\beta}}\right] = 1 - e^{-t};$$

$$G_2(x) = G_1(x) - A e^{-A}; \text{ etc.}$$

Out of there we deduce a non-recurrent explicit expression for (9):

$$G_i(x) = 1 - e^{-A}\left(1 + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \ldots + \frac{A^{i-1}}{(i-1)!}\right).$$

Let us find the ME of the i-th failure point:

$$M_i = \int_0^\infty (1 - G_i(x))dx =$$

$$= \int_0^\infty e^{-A}\left(1 + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \ldots + \frac{A^{i-1}}{(i-1)!}\right)dx;$$

the unknown ME is as follows:

$$M_{i+1} = C\left(\Gamma(\frac{1}{\beta}) + \Gamma(\frac{1}{\beta}+1) + \ldots + \frac{1}{i!}\Gamma(\frac{1}{\beta}+i)\right) =$$

$$= C\left(\Gamma(\frac{1}{\beta}) + \sum_{j=1}^{i}\frac{1}{j!}\Gamma(\frac{1}{\beta}+j)\right); \quad (15)$$

where C = const; G is a gamma function.

For $j > 3$ according to the properties of the gamma function the following is correct:

$$\frac{1}{j}\Gamma(\frac{1}{\beta}+j) > \frac{1}{j}\Gamma(j) = \frac{1}{j};$$

$$\sum_{j=1}^{i}\frac{1}{j!}\Gamma(\frac{1}{\beta}+j) > \sum_{j=1}^{i}\frac{1}{j} \to \infty;\ i \to \infty.$$

Therefore, the sequence (15) diverges, in other words:

$$\lim_{i\to\infty} M_i = \lim_{i\to\infty} M(S_i \mid q=1) = \lim_{i\to\infty} \sum_{j=1}^{i} M(X_i | q=1) = \infty. \quad (16)$$

I.e. for the specific case $q = 1$, the sequence of the MEs of failure points diverges under Weibull distribution for TFF and Kijima-1 and Kijima-2 models.

5.2 Divergence of failure points for incomplete recovery

Taking into consideration (2) and (3), for Kijima models the following is correct: the virtual age increases monotonously as the parameter q increases, other things equal for model Kijima-1 we have:

$$Mv_n = q\sum_{i=1}^{n} MX_i.$$

In general, if $q_1 < q_2$, then:

$$v_n(q_1) < v_n(q_2), \quad (17)$$

$$M(v_n(q_1)) < M(v_n(q_2)).$$

Next, we examine the conditional distribution function (1):

$$F_i(x \mid v_{i-1}) = \frac{F(x+v_{i-1}) - F(v_{i-1})}{1 - F(v_{i-1})}.$$

Under Weibull distribution (4) for TFF, the function changes to:

$$F_i(x \mid v_{i-1}) = 1 - \exp(\lambda({v_{i-1}}^{\beta} - (x+v_{i-1})^{\beta}).$$

Under fixed x this function monotonously increases as v increases, in other words, if $v^{*}_{i\text{-}1} < v^{**}_{i\text{-}1}$, then:

$$F_i(x \mid v^{*}_{i-1}) < F_i(x \mid v^{**}_{i-1}).$$

Under (17), if $q_1 < q_2$, then the following inequation is correct:

$$F_i(x \mid v_{i-1}(q_1)) < F_i(x \mid v_{i-1}(q_2)).$$

Whereas:

$$M\left(X_i \mid v_{i-1}(q)\right) = \int_0^{\infty} \left\{1 - F_i\left[x \mid v_{i-1}(q)\right]\right\} dx,$$

it follows whence:

$$M\left(X_i \mid v_{i-1}(q_1)\right) > M\left(X_i \mid v_{i-1}(q_2)\right),$$

$$MM\left(X_i \mid v_{i-1}(q_1)\right) > MM\left(X_i \mid v_{i-1}(q_2)\right),$$

out of which based on the ME property we get the inequation for the unconditional mean:

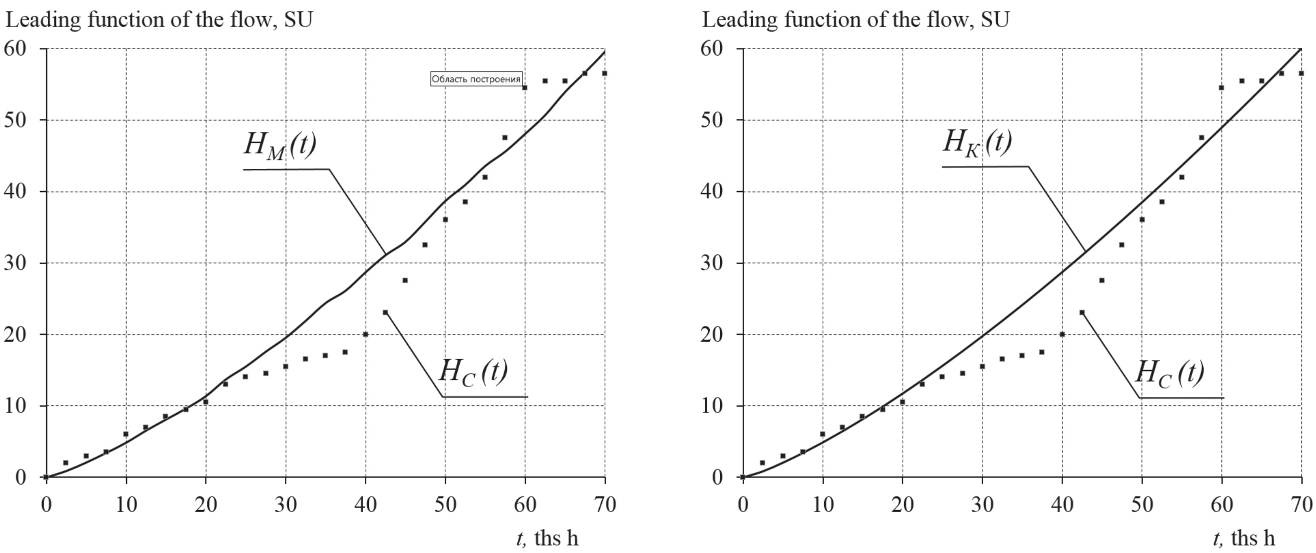
$$M\left(X_i(q_1)\right) > M\left(X_i(q_2)\right). \quad (18)$$

Subject to (16) and (18), we come to the following conclusion: if this sequence diverges under $q = 1$, then this sequence is also diverging under $q \in [0; 1]$ other things equal. Therefore, the sequence of the MEs of time points of the i^{th} failure of the Kijima-1 and Kijima-2 models and Weibull distribution for TFF **diverge** at $q \in [0; 1]$.

6. Example of calculation

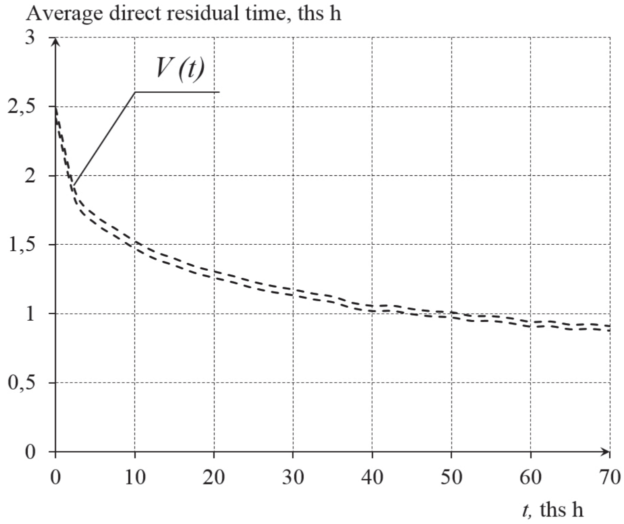
Let us try out the above examined models with real data. For that purpose let us use the information on the failures of information collection and processing devices operating as part of standard equipment of nuclear power plants. For two selected devices over the time T≈8Ч10^4 h, 121 failure was registered. The value of LLT (6), parameter estimates and confidential intervals for probability belief 0.95 are given in table 1:

Given the LLT values, the Kijima-1 model should be chosen for further research, as it has the highest LLT out of the compared models. Large confidential intervals of estimates are explained by the insufficient amount of input data. Point estimation of LLT values is given in Figure 2 with the following designations:

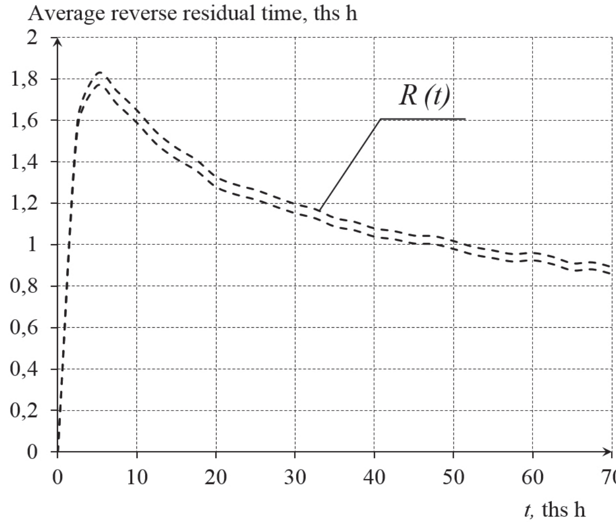


a) b)

Fig. 2. Estimation of LLT by means of a) statistical testing and b) sum total method



a)



b)

Fig. 3. Estimation of the average a) direct and b) reverse residual times

Table 1. Model parameters evaluation

Model	ln L	λ	β	q	Δλ	Δβ	Δq
Kijima 1	-135.353	0.264	1.345	0.384	0.186	0.196	0.574
Kijima 2	-135.707	0.244	1.293	1.000	0.192	0.172	0.043

$H_C(t)$ is the graduated empirical LLT:

$$H_C(t) = i,\ i : T_i \le t \le T_{i+1};$$

$H_M(t)$ is the LLT for Kijima-1 model by means of statistical testing using formula (11);

$H_K(t)$ is the LLT for Kijima-1 model by means of the sum total method using formulas (7) and (9);

Figure 2 shows that the compared methods yield practically identical results. LLT estimates $H_M(t)$ and $H_K(t)$ correspond well with the experimental data, the function $H_C(t)$, and are suitable for forecasting.

Interval estimations of the residual time are given in figure 3 with ehe following designation:

$V(t)$ is the ADRT estimate for Kijima-1 model using formulas (12) and (14);

$R(t)$ is the ARRT estimate for Kijima-1 model using formulas (13) and (14);

As the graph shows, the ADRT and ARRT estimates constantly decrease and within the observed period of time do not have an asymptotic limited value, which, among other things, indicates the absence of steady mode within the observed interval, progressive degradation of its characteristics and presence of incomplete recovery.

Conclusion

Kijima models allow taking into consideration incomplete recoveries “worse than new but better than before the failure”. The article analyses the method of obtaining interval estimations of Kijima models parameters using Weibull distribution for time to first failure. The negative relation of the variance value and the number of observed elements is proven. The divergence of sequence of mathematical expectations of time point of the i-th failure for Kijima models is proven.

A method of point estimation of the leading function of the stream based on sum total calculation is suggested. Also, the authors suggest using the statistical test method for evaluation of the average direct time and reverse residual time for Kijima models. The paper puts forward evaluations of dependability characteristics based on operation data of information collection and processing devices as part of standard equipment of nuclear power plants.

References

1. **Wibowo W.** On approaches for repairable system analysis : Renewal Process, Nonhomogenous Poisson Process, General Renewal Process /W. Wibowo // Indonesia, Jurnal Industri, 2010 – Vol.9, №1 – P.60-66.

2. **Kaminsky M., Krivtsov V.** Use of the Monte Carlo method in the evaluation of the extended recovery process as part of failure data analysis during the warranty period // Reliability: Theory & Applications. – 2006. – Issue No. 1 – P. 32 – 34.

3. **Mettas A., Zhao W.** Modeling and Analysis of Repairable Systems with General Repair // Reliability and Maintainability Symposium, 2005 – P.176-182.

4. **Chumakov I.A., Antonov A.V.** Evaluation of dependability characteristics on the assumption of incomplete recovery // Dependability. – 2014. – Issue No. 1 (48). – P. 3 11.

5. **Guo H., Liao H., Pulido J.** Failure Process Modeling For Systems With General Repairs // MMR 2011. International Conference on Mathematical Methods in Reliability, 2011.

6. **Antonov A.V.** System analysis. Textbook for higher educational institutions / – Moscow: Vyshaya shkola, 2004. – 454 p.

7. **Beihelt F., Franken P.,** Reliability and maintenance / – Moscow: Radio i sviaz, 1988. – 357 p.

8. **Bakhvalov N.S., Zhidkov N.P., Kobelkov G.M.** Numerical techniques / – Moscow: Nauka, 2003. – 635 p.

9. **Chepurko V.A., Chepurko S.V.** Non-uniform streams models in the renewal theory / – Obninsk: NRNU MEPhI, 2012. – 162 p.