

- 3. // - 1992. - 3,4 - . 13-20.
- 4. ,, . - , 1996 - 276 .
- 5. ,, - 1997. - 1,2 - . 35-38.

6. *A.Nikov, T.Georgiev. A fuzzy neural network and its MatLab Simulation.* Proc. of ITISS 21-st Intern. Conference Technology on Inform // Technology Interfaces. - 1999. - P. 413-418.

13.06.01.

fcoVC 3-  
519.24:519.652.3

.. , .. , 0.0.

$\gamma(\cdot; \odot) = \{ \cdot : (\cdot) \}$

$$\langle \Omega_n, A, P_n \rangle, \text{ де } \Omega_n \in \Omega, \Omega_n = \{ x_i; i = \overline{1, n} \} -$$

/'', ( )

( )

( - )

$X - ,, = 0, / -1>, \sim \mathcal{L} ,$

= \* + ( - ) + /2,

де  $h$  – крок розбиття:  $h = \frac{x_{\max} - x_{\min}}{r}$ ;  $x_{\min}$ ,  $x_{\max}$  – мінімальне та максимальне значення варіаційного ряду;  $r$  – кількість варіантів.

Для варіаційного ряду, розбитого на класи, обчислені частоти, випадковості та емпірична функція розподілу:  $n = \{n_i, i = \overline{0, r-1}\}$ ,  $f = \{f_i, i = \overline{0, r-1}\}$ ,  $F_n = \{Fn_i, i = \overline{0, r-1}\}$  відповідно. Необхідно знайти неперервне наближення функції щільності розподілу ймовірності реалізацій випадкової величини  $f(x)$ .

За таке наближення виберемо локальний поліноміальний сплайн на основі  $B$ -сплайнів другого порядку, близький до інтерполяційного у середньому. Тоді наближення функції  $f(x)$  у вигляді локального сплайна буде таке [1]:

$$S_{2,0}(p, t) = \sum_{i \in Z} p_i B_{2,h}(t - ih), \quad (1)$$

де

$$B_{2,h}(t) = \begin{cases} 0, & |t| \geq 3h/2; \\ \frac{(3 + 2t/h)^2}{8}, & t \in [-3h/2, -h/2]; \\ \frac{3}{4} - \frac{(2t/h)^2}{4}, & t \in [-h/2, h/2]; \\ \frac{(3 - 2t/h)^2}{8}, & t \in [h/2, 3h/2] \end{cases}$$

або у розгорнутому вигляді:

$$S_{2,0}(p, y) = \frac{1}{8} \left( (p_{i-1} + 6p_i + p_{i+1}) + (-2p_{i-1} + 2p_{i+1})y + (p_{i-1} - 2p_i + p_{i+1})y^2 \right), \quad (2)$$

де  $p_i = f_i$ ;  $y = \frac{2}{h}(t - x_i) - 1, |y| \leq 1$ ;  $i = \left[ \frac{t - x_{\min}}{h} \right] + 1$ ;  $t \in [x_{\min}, x_{\max}]$ .

Як невизначені у ході експерименту значення  $f_{-1}$ ,  $f_r$  цілком достатньо взяти величини  $f_{-1} = f_0$ ,  $f_r = f_{r-1}$ .

Отже, для будь-якого  $t \in [x_{\min}, x_{\max}]$  за допомогою описаного підходу на основі локальних поліноміальних сплайнів можна обчислити функцію щільності розподілу ймовірності реалізацій випадкової величини  $f(t)$  за таким алгоритмом:

1) визначається клас варіаційного ряду, до якого належить точка  $t$ :  $i = \left[ \frac{t - x_{\min}}{h} \right] + 1$ ;

2) обчислюється  $y = \frac{2}{h}(t - x_i) - 1$ ;

3) обчислюється значення функції щільності в точці  $t$  у вигляді (2).

Для відтворення функції розподілу ймовірності реалізацій випадкової величини пропонується аналогічний підхід. За наближення виберемо локальний поліноміальний сплайн на основі  $B$ -сплайнів третього порядку, близький до інтерполяційного в середньому. Його використання для відтворення функції розподілу впливає з властивостей  $B$ -сплайнів та функції розподілу:

$$B_{k,h}(t) = \frac{1}{h} \int_{t-h/2}^{t+h/2} B_{k-1,h}(t) dt, \quad k = 1, 2, \dots, \quad F(t) = \frac{1}{h} \int_{x_i}^{x_{i+1}} f(t) dt$$

Отже, для відтворення функції розподілу ймовірності реалізацій випадкової величини будемо шукати неперервне наближення функції  $F(x)$  у вигляді локального сплайна:

$$S_{3,0}(p, t) = \sum_{i \in Z} p_i B_{3,h}(t - ih), \quad (3)$$

$$\text{де } B_{3,h}(t) = \frac{1}{48} \begin{cases} 0 & (-\infty, -2h]; \\ (4 + 2t/h)^3 & [-2h, -h]; \\ -3(2t/h)^3 - 12(2t/h)^2 + 32 & [-h, 0]; \\ 3(2t/h)^3 - 12(2t/h)^2 + 32 & [0, h]; \\ (4 - 2t/h)^3 & [h, 2h]; \\ 0 & [2h, \infty) \end{cases}$$

або у розгорнутому вигляді:

$$\begin{aligned} S_{3,0}(p, y) = & \frac{1}{48}(-p_{i-1} + 3p_i - 3p_{i+1} + p_{i+2})y^3 + \frac{1}{16}(p_{i-1} - p_i - p_{i+1} + p_{i+2})y^2 + \\ & + \frac{1}{16}(-p_{i-1} - 5p_i + 5p_{i+1} + p_{i+2})y + \frac{1}{48}(p_{i-1} + 23p_i + 23p_{i+1} + p_{i+2}), \end{aligned} \quad (4)$$

$$\text{де } p_i = Fn_i; \quad y = \frac{2}{h}(t - x_i) - 1, |y| \leq 1; \quad i = \left\lfloor \frac{t - x_{\min}}{h} \right\rfloor + 1; \quad t \in [x_{\min}, x_{\max}].$$

Як невизначені у ході експерименту значення  $Fn_{-1}, F_r, F_{r+1}$  цілком достатньо взяти величини  $Fn_{-1} = 0, Fn_r = 1, Fn_{r+1} = 1$ .

Отже, для будь-якого  $t \in [x_{\min}, x_{\max}]$  можна обчислити функцію розподілу ймовірності реалізацій випадкової величини  $F(t)$  за аналогічним алгоритмом:

$$1) \text{ обчислюється } i = \left\lfloor \frac{t - x_{\min}}{h} \right\rfloor + 1;$$

$$2) \text{ обчислюється } y = \frac{2}{h}(t - x_i) - 1;$$

3) визначається значення функції розподілу в точці  $t$  за виразом (4).

Вибір сплайнів (1), (3) обумовлений високими апроксимаційними властивостями [2]. Простота реалізації забезпечує швидкодію та ефективність обчислювальної схеми з використанням сплайнів (2), (4).

Поряд з обчисленням значень функцій розподілів досить часто доводиться розв'язувати обернену задачу – обчислення квантилів розподілу за даною ймовірністю.

Для знаходження квантилів функції розподілу, одержаної за допомогою застосування локального поліноміального сплайна на основі  $B$ -сплайнів третього порядку, пропонується така ітераційна процедура. Нехай маємо варіаційний ряд  $X = \{x_i, i = \overline{0, r-1}\}$  з неперервною функцією розподілу ймовірностей  $F(x_i) \in C^3$ . Для заданої ймовірності  $P$  квантиль знаходиться за ітераційним алгоритмом:

1) покладемо  $x^* = \frac{x_{\min} + x_{\max}}{2}$ , де  $[x_{\min}, x_{\max}]$  – інтервал, на якому відшукується значення квантиля (спочатку  $x_{\min}, x_{\max}$  найменше та найбільше значення варіаційного ряду відповідно);

2) обчислюється  $S_{30}(x^*)$  за виразом (4);

3) якщо  $|S_{30}(x_*) - P| < \epsilon$ , де  $\epsilon$  – задана точність, то процес завершується, а за квантиль обирається значення  $x^*$ , у противному разі переходимо на крок 4;

4) якщо  $S_{30}(x_0) > P$ , то значення  $x_{\max}$  замінюється на  $x^*$ , тобто далі обирається інтервал  $[x_{\min}, x^*]$ , інакше проводиться заміна  $x_{\min}$  на  $x^*$ , тобто далі розглядається інтервал  $[x^*, x_{\max}]$  і переходимо на крок 1.

Для перевірки якості наведених алгоритмів було проведено серії експериментів з генерування масивів, розподілених за нормальним законом розподілу з наступним їх відтворен-

ням. Вибір нормального закону розподілу пояснюється тим, що в основу практично всієї класичної теорії ймовірностей та прикладного статистичного аналізу покладено цей розподіл, який є межевою формою багаточисельних розподілів. Експеримент з генерації масивів обсягу 7000 реалізацій випадкових величин, розподілених за нормальним законом розподілу, проводився для  $m = 0$ ,  $\sigma = 1$ .

Імітація нормально розподілених випадкових величин здійснюється неоднозначно у залежності від реалізованого методу. Задовільну достовірність узгодження з нормальним законом дає алгоритм генерування випадкових чисел, заснований на властивості, що множення випадкових величин, розподілених за законом Релея з параметром  $\sigma$  і за законом арксинуса з параметрами  $m = 0$  і  $\lambda = 0,5$ , є випадковою величиною, розподіленою за нормальним законом з параметрами  $(0, \sigma)$  [3].

За таким способом для двох рівномірно розподілених випадкових чисел на  $[0,1)$   $R_1, R_2$  отримують із співвідношень  $X_1 = \sqrt{-2 \ln R_1} \sin 2\pi R_2$ ,  $X_2 = \sqrt{-2 \ln R_1} \cos 2\pi R_2$  два випадкових числа  $X_1, X_2$ , що мають нормальний розподіл  $N(0,1)$ .

Далі для згенерованого масиву відбувається побудова варіаційного ряду, розбиття його на класи, а також обчислення частот, випадковостей та емпіричної функції розподілу.

За наведеними алгоритмами відбувається відтворення функції щільності та функції розподілу за нормальним законом розподілу та обчислення квантилів.

У ході експерименту за згенерованими даними проводиться відтворення функції розподілу за нормальним законом розподілу та обчислення квантилів за допомогою класичних апроксимаційних формул [4], згідно з якими функція нормального розподілу має вигляд:

$$F(x) = N(x, m, \sigma) = \frac{1}{\sqrt{2\pi\sigma}} \int_{-\infty}^x \exp\left(-\frac{(y-m)^2}{2\sigma^2}\right) dy = \Phi\left(\frac{x-m}{\sigma}\right) = \Phi(u),$$

де  $\Phi(\bullet)$  – функція Лапласа.

Для функції розподілу, що визначається через  $\Phi(u)$ , має місце асимптотичний вираз:

$$\Phi(u) = 1 - \frac{1}{u\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right) \left(1 - \frac{1}{u^2} + \frac{1 \cdot 3}{u^4} - \dots - \frac{(-1)^n 1 \cdot 3 \dots (2n-1)}{u^{2n}}\right).$$

За рахунок перетворень остаточно одержують:

$$\Phi(u) = 1 - \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right) (b_1 t + b_2 t^2 + b_3 t^3 + b_4 t^4 + b_5 t^5) + \varepsilon(u),$$

де  $b_1 = 0,319381530$ ;  $b_2 = -0,356563702$ ;  $b_3 = 1,7814779377$ ;

$b_4 = -1,821255978$ ;  $b_5 = 1,330274429$ ;  $\rho = 0,2316419$ ;  $|\varepsilon(u)| \leq 7,8 \cdot 10^{-8}$ .

Квантилі нормально розподіленої випадкової величини отримують з рівняння

$$\frac{1}{\sqrt{2\pi}} \int_{u_p}^{\infty} \exp\left(-\frac{u^2}{2}\right) du = p \quad \text{або} \quad u_p = \Phi^{-1}(1-p).$$

Значення  $u_p$  може бути обчислено за такою апроксимаційною формулою:

$$u_p = t - \frac{c_0 + c_1 t + c_2 t^2}{1 + d_1 t + d_2 t^2 + d_3 t^3} + \varepsilon_p,$$

де  $t = \sqrt{\ln \frac{1}{p^2}}$ ;  $c_0 = 2,515517$ ;  $c_1 = 0,802863$ ;  $c_2 = 0,010328$ ;

$d_1 = 1,432788$ ;  $d_2 = 0,189269$ ;  $d_3 = 0,001208$ ;  $|\varepsilon_p| \leq 4,5 \cdot 10^{-4}$ .

Аналіз якості описаних методів проводився на основі порівняння квантилів, обчислених за допомогою запропонованих алгоритмів на основі локальних поліноміальних сплайнів з квантилями, наведеними у статистичних таблицях, та квантилями, отриманими за допомогою апроксимаційних формул шляхом обчислення відносних похибок:

$$\delta_1 = \frac{u_p - u_\alpha}{u_\alpha} 100, \quad \delta_2 = \frac{u^* - u_\alpha}{u_\alpha} 100, \quad \delta_3 = \frac{u^* - u_p}{u_p} 100,$$

де  $u_p$  – квантиль, обчислений за допомогою апроксимаційних формул;  $u_\alpha$  – квантиль, наведений у статистичних таблицях;  $u^*$  – квантиль, обчислений за допомогою описаного методу на основі локальних поліноміальних сплайнів.

Результати експерименту наводяться у табл. 1,2 та на рис. 1,2.

Таблиця 1

Порівняння результатів роботи методів для масиву даних, згенерованого за нормальним розподілом (кількість елементів = 7000, кількість класів = 17)

Ймовірність	$u_\alpha$	$u_p$	$\delta_1, \%$	$u^*$	$\delta_2, \%$	$\delta_3, \%$
0,005	-2,57	-2,5759	0,2307	-2,5880	0,7012	0,4694
0,025	-1,96	-1,9602	0,0077	-2,0210	3,1106	3,1027
0,050	-1,65	-1,6450	0,0001	-1,6921	2,8619	2,8619
0,100	-1,28	-1,2816	0,1213	-1,3032	1,8158	1,6925
0,250	-0,68	-0,6741	0,1381	-0,6892	2,1045	2,2456
0,500	0,00	0,0001	0,0069	0,0123	1,2320	1,2250
0,750	0,68	0,6668	1,2082	0,7025	4,0746	5,3473
0,900	1,28	1,2410	3,4215	1,3230	2,9588	6,6063
0,950	1,65	1,5559	5,7018	1,6892	2,3758	8,5660
0,975	1,96	1,7995	8,1880	1,9776	0,8960	9,8941
0,995	2,58	2,1664	16,0301	2,5705	0,3668	18,6535

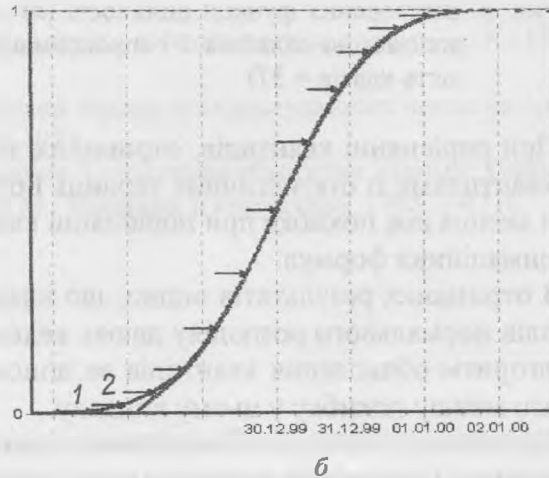
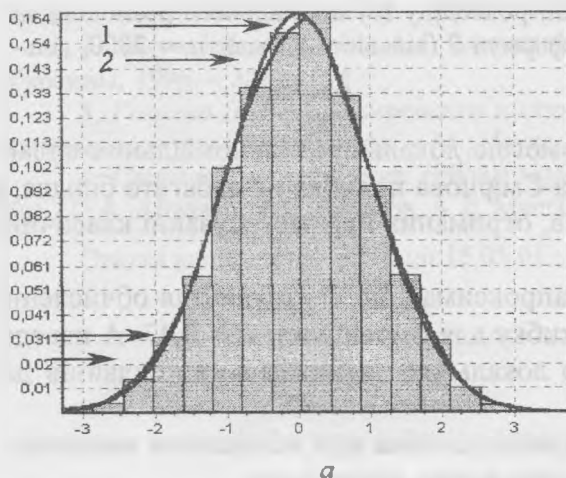


Рис. 1. Відтворення функції щільності (а) і функції розподілу (б) нормального розподілу за допомогою сплайнів 1 і апроксимаційних формул 2 (кількість елементів = 7000, кількість класів = 17)



можемо віддавати перевагу вказаним алгоритмам. Далі наводиться робота описаних алгоритмів на прикладі вибірки, що містить інформацію про час напрацювання на відмову приладу ЕУП-53 літака. Результати відтворення у вигляді стійких квантилів зведено до табл. 3, відтворені функції щільності та розподілу наведені на рис.3.

Ймовірність	0,005	0,025	0,050	0,100	0,250	0,500	0,750	0,900	0,950	0,975	0,995
$u^*$	1035,29	1102,82	1570,59	1825,69	2919,28	4974,72	7232,99	8792,88	9241,39	9399,26	9503,81

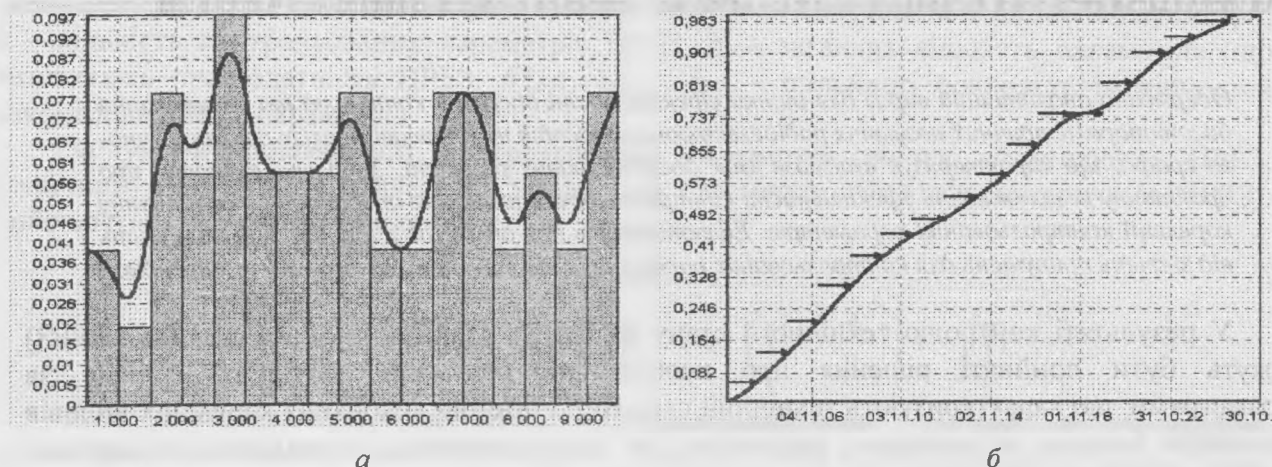


Рис. 3. Відтворені функції щільності (а) і функції розподілу (б) часу наробітку на відмову приладу ЕУП-53 літака

Аналіз отриманих результатів підтверджує обґрунтованість пропозиції щодо використання зазначеного підходу непараметричної оцінки функції розподілу ймовірностей при розв'язку практичних задач.

#### Список літератури

1. Лигун А.А., Кармазина В.В. Восстановление эмпирической функции плотности распределения локальными сплайнами // Современные вопросы теории приближения и комплексного анализа. – К.: Ин-т математики АН УССР. – 1990. – С. 78 – 88.
2. Лигун А.А., Шумейко А.А. Асимптотические методы восстановления кривых. – К.: ИМ НАН Украины, 1996. – 358 с.
3. Голенко Д.И. Моделирование и статистический анализ псевдослучайных чисел на электронных вычислительных машинах. – М.: Наука, 1965. – 107 с.
4. Прикладная статистика: основы моделирования и первичная обработка данных: Справ. пособие / С.А. Айвазян, И.С. Енюков, А.Д. Мешалкин. – М.: Финансы и статистика, 1983. – 471с.

Стаття надійшла до редакції 15.05.01.