УДК 519.2

А. Левченко
Академія сухопутних військ
імені гетьмана Петра Сагайдачного

ПОНЯТТЯ ДОСТОВІРНОСТІ МАТЕМАТИЧНИХ МОДЕЛЕЙ ВИПАДКОВИХ ПРОЦЕСІВ КЕРУВАННЯ СТАНОМ СКЛАДНИХ ОБ'ЄКТІВ

За результатами аналізу функціонування складних систем розглянуто специфіку відносин між об'єктом дії системи та його моделлю. Пропонується реалізація поняття достовірності модельної оцінки об'єкта. Для введеного поняття наведено кількісні співвідношення, що дають змогу оцінювати числові значення достовірності моделей за кінцевим числом реалізацій з урахуванням випадкової і систематичної складової статистичних відомостей.

On the results of analysis of functioning of the difficult systems the specific of relations is considered between the object of action of the system and his model. Realization of concept of authenticity of model estimation of object is offered. For the entered concept quantitative correlations which allow to estimate the numerical values of authenticity of models after the eventual number of realization taking into account the casual and systematic constituent of statistical information are resulted.

Постановка проблеми. Аналіз останніх досліджень та публікацій

У [1] під час аналізу функціонування системи забезпечення експлуатації складних технологічних систем обслуговування визначено завдання і принципи дії кожної підсистеми як специфічні функціональні ланки. Так, завдання підсистеми контролю полягає у визначенні стану і властивостей об'єкта експлуатації, що подається у вигляді цифрової вимірювальної інформації. Числова вимірювальна інформація, яку здобувають у результаті операцій функціонування підсистеми контролю (метрологічного забезпечення), використовується в інформаційних системах підтримки прийняття рішень про введення керуючого впливу, для управління процесом. Модель процесу є основою таких систем підтримки прийняття рішень.

У загальному випадку аналітична модель, що відображає в підсистемі прийняття рішень реальний об'єкт або ситуацію, може бути представлена у вигляді крапки у фазовому просторі фізичних величин (характеристик, технічних параметрів тощо); координатні осі цього фазового простору є відображенням визначених реально існуючих властивостей об'єкта [1].

Реальні об'єкти впливу системи знаходяться в множинних зв'язках і взаємодіях з об'єктами матеріального світу, не введеними до системи, внаслідок чого остання зазнає таких змін, які не можна чітко врахувати при побудові сукупності моделей конкретної системи.

Цей фактор є об'єктивною причиною випадковості подій матеріального світу і основою понять "випадкова величина", "випадковий процес". У загальному випадку всі фізичні величини, що вимірюються під час наукових досліджень, виробництва, а також в процесі експлуатації різних технічних об'єктів, мають випадковий характер, що визначається їхньою мінливістю під впливом сукупності зовнішніх і внутрішніх чинників, які не піддаються обліку.

Отже, визначаючи вимірювання як "знаходження значення фізичної величини дослідним шляхом..." [2], треба пам'ятати, що вимірювана величина, будучи випадковою, не має конкретного точного значення у власному сенсі цього слова, а може бути представлена безліччю (дискретних або континуальних) значень із зазначенням ймовірності кожного з них. При інтервальному оцінюванні випадкової величини X зазначають числове значення ймовірності $P[X \in (a,b)]$, з якою випадкова величина потрапляє у фіксований інтервал значень (a,b).

Експериментально значення ймовірності $P[X \in (a,b)]$ шукають шляхом багатократного спостереження реалізацій випадкової величини X і подальшого узагальнення результатів цих спостережень.

Що особливо слід зазначити: як випливає з визначення ймовірності [3], її точне значення можна оцінити як межу за нескінченно великої кількості спостережень. Оскільки на практиці нескінченна кількість спостережень не реалізовується, точне значення ймовірності P визначити неможливо, тому доводиться оперувати оцінкою $\tilde{P}$, більшою чи меншою мірою наближеною до її реального значення.

Крім того, в експериментальній процедурі оцінювання спостерігаються не самі реалізації $x_1, x_2, ..., x_n$ випадкової величини X, а результати $x_1, x_2, ..., x_n$ вимірювань цієї величини.

Оскільки засобам і методам вимірювань властиві похибки, реальні значення x_i величини X у момент і-го спостереження в загальному випадку не збігаються з отриманими результатами $x_з$, внаслідок чого навіть за великої кількості n спостережень значення ймовірності P неможливо визначити точно.

Формулювання цілі статті

Обмеженість кількості спостережень і наявність похибок вимірювань приводять до невизначеності оцінок величин, що описують реальні об'єкти. Ця невизначеність разом з об'єктивно існуючою випадковістю самих оцінюваних величин обумовлює неможливість точного прогнозу результатів реалізації системою вироблених управляючих рішень, що знаходить віддзеркалення в імовірнісному характері оцінок ефективності систем, їх здатності виконувати вказані завдання.

Отже, разом з поняттям ймовірності події, що визначає міру його об'єктивної можливості [4], доцільно вводити поняття достовірності оцінки цієї події, що характеризує міру відповідності між реальною подією і його моделлю, що є метою роботи.

Виклад основного матеріалу

Достовірність оцінки залежить від двох основних чинників: кількості спостережень при вимірюванні випадкової величини і похибки вимірювань.

Для числового вимірювального контролю відносно випадку оцінювання випадкової величини у вигляді параметра об'єкта, розглядаємо ймовірність події A, що полягає у потраплянні X в інтервал полів допусків (a,b) на параметр.

Нехай оцінювання $P(A)$ проводиться за результатами спостережень $x_1, x_2, ..., x_n$; тоді вважаємо, що достовірність $v[\tilde{P}(a,b)]$ оцінки $\tilde{P}(A)$ шуканої ймовірності за визначенням чисельно дорівнює імовірності $P(x_1, x_2, ..., x_n)$ результатів, що спостерігалися за умови реалізації події A.

Оскільки $\tilde{P}(A)$ – континуальна величина на інтервалі

$$0 \le \tilde{P}(A) \le 1,$$

у [2] вперше було запропоновано введення функції достовірності $v(\tilde{P})$ оцінки, визначивши її співвідношенням

$$J(\tilde{P}) = D \cdot \lim_{\Delta \tilde{P} \to 0} \frac{V(\tilde{P})}{\Delta P},$$

де D – постійний нормуючий множник.

Існує аналогія між функцією достовірності оцінки і функцією розподілу ймовірності випадкової величини, проте кожна з них відображає специфічні сторони взаємодії системи з матеріальними об'єктами, а саме: перша характеризує невизначеність знакового відображення і визначається на безлічі моделей, друга ж описує мінливість поведінки самого матеріального об'єкта, що породжується фізичними причинами. Вид функції достовірності показаний на рис. 1.

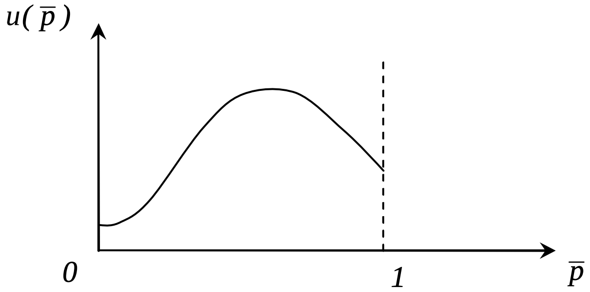

Рис. 1. Загальний вигляд функції достовірності інтервальної оцінки випадкової величини

Саме тому говорять про достовірність моделі, а не про ймовірність здійснення події. Отримавши при вимірюванні ряд спостережень $x_1, x_2, ..., x_n$ (показників засобів вимірювань), слід розмірковують про вимірювану величину тільки в сенсі відшукання для кожного фіксованого інтервалу її можливих значень відповідної функції достовірності $J[\tilde{P}(a,b)]$. Форма самої функції відображає якість методики виконання технологічних операцій вимірювання, тобто метрологічні властивості засобів вимірювань, кількість спостережень тощо.

Так, при кількості спостережень $n \to \infty$ і методиці виконання вимірювань, позбавленої методичної складової похибок:

$$J(\tilde{P}) = d(\tilde{P} - P),$$

де P – реальне значення ймовірності.

Відомо, що випадкова величина $\tilde{O}$ описується функцією щільності розподілу ймовірності $f(x)$, що отримується з безлічі інтервальних оцінок шляхом граничного переходу [2, 4, 5] у вигляді:

$$f(x) = \lim_{D\tilde{P} \to 0} \frac{P[x \in (x, Dx)]}{Dx},$$

де Dx – інтервал, що оточує фіксоване значення x.

У реальних вимірюваннях з обмеженою кількістю спостережень і точністю, для кожного x отримуємо безліч оцінок $f(x)$, кожна з яких характеризується числовим значенням функції достовірності $J[f(x)]$ і розподіл $f(x)$ набуває характеру випадкової функції. Графічна форма цієї функції схожа на функцію дрейфу числового параметра з двостороннім полем допуску.

Значення достовірності оцінки функції $f(x)$ для фіксованої точки x характеризується "щільністю" в точці з координатами $(x, \tilde{f}(x))$.

Тільки за абсолютно точних вимірювань і нескінченно великої кількості спостережень можливо отримати реальне значення розподілу випадкової величини, як зображено на рис. 2.

Розглянемо загальний випадок інтервального оцінювання, в якому потрібно визначити функцію достовірності $J[\tilde{P}(a,b)]$ оцінки $\tilde{P}$ ймовірності потрапляння випадкової величини X в заданий інтервал (a,b) при вимірюванні, що має похибка D. Результати спостережень можна подати у вигляді

$$x_1 = X_1 + \overset{0}{D}_1 + D_{C1},$$

$$x_2 = X_2 + \overset{0}{D}_2 + D_{C2},$$

$$\ldots\ldots\ldots\ldots\ldots\ldots,$$

$$x_i = X_i + \overset{0}{D}_i + D_{Ci},$$

$$\ldots\ldots\ldots\ldots\ldots\ldots,$$

$$x_n = X_n + \overset{0}{D}_n + D_{Cn},$$

де X_1 X_i – значення випадкової величини X в момент і-го спостереження; $\overset{o}{D}_i$ – значення випадкової складової похибки вимірювання; D_{C_i}, – значення систематичної складової похибки вимірювання.

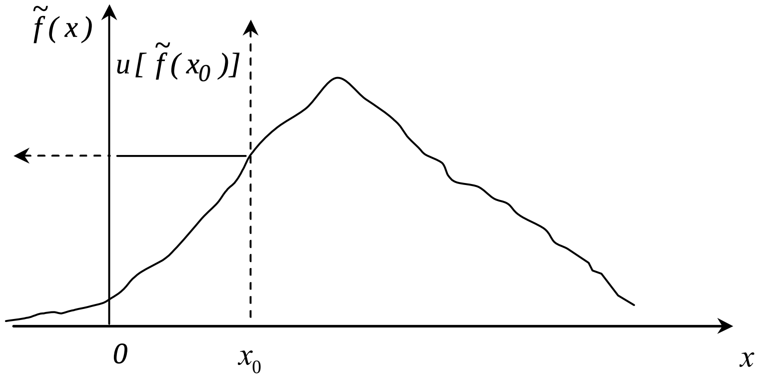

Рис. 2. Оцінка $\tilde{f}(x)$, що збігається з реальним розподілом $f(x)$

Вважаючи систематичну складову похибки у цій серії спостережень незмінною, отримаємо:

$$D_{Ci} = D_{Cj} = D_C,\ i \neq j = 1,...,n.$$

Систематична складова D_C, будучи точно невідомою, є невизначеною величиною і тому може бути описана функцією достовірності $v(\overline{D}_C)$, а сумарна похибка

$$D = \overset{0}{D} + D$$

опишеться умовною функцією розподілу

$$f_D(x/x, D_c) = f_u(x - D_c/x), \qquad (1)$$

де $f_u(x)$ – функція розподілу ймовірностей випадкової складової D похибки вимірювання.

Через наявність похибки вимірювання D точне число n^* потраплянь значень випадкової величини в заданий інтервал при реалізаціях (спостереженнях) залишається невідомим і може, взагалі кажучи, набувати значення

$$0 \leq n^* \leq n.$$

При цьому будь-яке значення n^* можна оцінити з певною достовірністю $V(\tilde{m}^n)$.

Тоді, за аналогією з формулою повної ймовірності, можна визначити

$$J(\tilde{p}) = \sum_{n^*=0}^{n} J(\tilde{p}/\tilde{n}^*) * V(\tilde{n}^*). \qquad (2)$$

Визначимо умовну функцію достовірності $J(\tilde{p}/\tilde{n}^*)$, припустивши, що число n^* набуде деякого фіксованого $\tilde{n}^*$ значення.

Відповідно введеному визначенню достовірності оцінки шукатимемо ймовірність рівно $\tilde{n}^*$ потраплянь випадкової величини X в інтервал (x, x^*) з n реалізацій (дослідів, спостережень).

Ця ймовірність визначиться співвідношенням [3]

$$P(\tilde{n}^*, n) = C_n^{\tilde{n}^*} \cdot \tilde{P}^{\tilde{n}^*} \cdot (1 - \tilde{P})^{n - \tilde{n}^*}, \qquad (3)$$

звідки

$$J(\tilde{P}, \tilde{n}^*, n) = D \cdot C_n^{\tilde{n}^*} \tilde{P}^{\tilde{n}^*} \cdot (1 - \tilde{P})^{n - \tilde{n}^*}. \qquad (4)$$

Значення нормувального множника D знайдемо з умови

$$\int_0^1 J(\tilde{P})d\tilde{P} = 1 .$$

Проінтегрувавши праву частину (4), після перетворень отримаємо

$$J(\tilde{P}/\tilde{n}^*) = (n+1)\cdot C_n^{\tilde{n}^*}\cdot \tilde{P}^{\tilde{n}^*}\cdot (1-\tilde{P})^{n-\tilde{n}^*} . \qquad (5)$$

Після підстановки виразу (5) в (2), маємо

$$J(\tilde{P}) = \sum_{n^*=0}^{n}(n+1)C_n^{\tilde{n}^*}\cdot \tilde{P}^{\tilde{n}^*}\cdot (1-\tilde{P})^{n-\tilde{n}^*}\cdot V(\tilde{n}^*) . \qquad (6)$$

Отриманий вираз (6) є формулою повної функції достовірності оцінки ймовірності P за наявності n спостережень.

Визначимо достовірність $V(\tilde{n}^*)$ оцінки $\tilde{n}^*$ за наявності результатів спостережень $x_1, x_2, ..., x_n$, для чого шукатимемо ймовірність $P(x_1, x_2, ..., x_n)$ при фіксованому значенні $n^* = \tilde{n}^*$.

Кожному спостереженню x_i відповідає реалізація X_i випадкової величини X. Сукупності n спостережень відповідають n реалізацій $X_1, X_2 ..., X_n$, кожна з яких може потрапляти або не потрапляти в інтервал (x', x'').

Припустімо, що спостереженню x_i відповідає реалізація X_i, що потрапляє в інтервал (x', x''), як показано на рис 3.

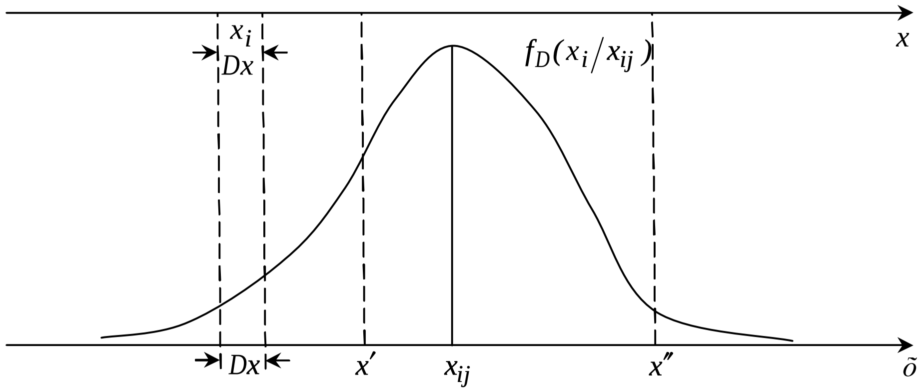

Рис. 3. До визначення оцінки $V[X_з \in (x', x'')]$

Якщо X_i набуде фіксованого значення X_{ij}, то ймовірність $P(x_i)$, як видно з рис. 3, становить

$$P(x_i/X_{ij}) = f(x_i/X_{ij})\cdot Dx = f_D(x_i/X_{ij})Dx .$$

Враховуючи (1), маємо

$$P(x_i/X_{ij}) = f_0(x_i - X_{ij} - D_c)\cdot Dx .$$

а для всіх можливих значень $X_i \in (x', x'')$ отримаємо

$$P(x_i/X_i \in (x', x'')) = \int_0^{\infty} f_0(x_i - D_c - x)\cdot dx . \qquad (7)$$

Перетворимо вираз (7), замінивши змінну інтегрування

$$P(x_i/X_i \in (x', x'')) = \int_{x_i - D_c - x''}^{x_l - D_c - x'} f_0(y)\cdot dy \qquad (8)$$

Очевидно, що

$$P(x_i / X_i \notin (x',x'')) = 1 - \int\limits_{x_i - D_c - x''}^{x_l - D_c - x'} f_0\,(y) \cdot dy\,. \qquad (9)$$

Отже, достовірність оцінки $V(\,X_i \in x',x''\,)$ при значенні, що спостерігається x_i і фіксованому значенні D_c становить величину

$$P(\,X_i \in (x',x'')/D_c\,) = \int\limits_{x_i - D_c - x''}^{x_l - D_c - x'} f_0\,(y) \cdot dy \qquad (10)$$

Припустімо, що при вимірюванні випадкової величини X отримано два результати x_1 і x_2, коли можливі наступні випадки:

$$A_1 \leftrightarrow (\,X_1 \in (x',x''))\cap(\,X_2 \in (x',x''))$$
$$A_2 \leftrightarrow (\,X_1 \in (x',x''))\cap(\,X_2 \notin (x',x''))$$
$$A_3 \leftrightarrow (\,X_1 \notin (x',x''))\cap(\,X_2 \in (x',x''))$$
$$A_4 \leftrightarrow (\,X_1 \notin (x',x''))\cap(\,X_2 \notin (x',x''))$$

при цьому

$$(\,n'' = 0\,) \leftrightarrow A_4$$
$$(\,n'' = 1\,) \leftrightarrow A_2 \cup A_3$$
$$(\,n'' = 2 = n\,) \leftrightarrow A_1$$

На підставі співвідношень (9) і (10) маємо

$$V(\tilde{n}^* = 0) = V(\,A_4\,) = (1 - \int\limits_{x_i - D_c - x''}^{x_l - D_c - x'} f_0(\,y\,)dy\,) \cdot (1 - \int\limits_{x_i - D_c - x''}^{x_l - D_c - x'} f_0(\,y\,)dy\,);$$

$$V(\tilde{n}^* = 1) = V(\,A_2\,) + V(\,A_3\,) = \int\limits_{x_i - D_c - x''}^{x_l - D_c - x'} f_0(\,y\,)dy \cdot (1 - \int\limits_{x_i - D_c - x''}^{x_l - D_c - x'} f_0(\,y\,)dy\,) +$$

(11)

$$+(1 - \int\limits_{x_i - D_c - x''}^{x_l - D_c - x'} f_0(\,y\,)dy\,) \cdot \int\limits_{x_i - D_c - x''}^{x_l - D_c - x'} f_0(\,y\,)dy;$$

$$V(\tilde{n}^* = 2) = V(\,A_1\,) = \int\limits_{x_i - D_c - x''}^{x_l - D_c - x'} f_0(\,y\,)dy \cdot \int\limits_{x_i - D_c - x''}^{x_l - D_c - x'} f_0(\,y\,)dy\,)$$

Аналогічно можна отримати вирази для достовірності оцінки $\tilde{n}^*$ за будь-якої кількості n спостережень.

Вираз (11), а отже, і формула (6) отримані для деякого фіксованого значення систематичної складової похибки D_c, що зберігається незмінним для всієї серії спостереження.

Невизначену величину D_c можна оцінити і описати функцією достовірності $u(D_c\,)$, визначуваною на основі наявної інформації про можливі або спостережені значення D_c [7,8].

У загальному випадку

$$-\infty \le \tilde{D}_c \le \infty,$$

тоді формула повної функції достовірності оцінки $\tilde{P}$ набуває вигляду

$$J(\tilde{P}\,) = \int\limits_{-\infty}^{\infty} \sum_{n^*=0}^{n} (n+1) C_n^{\tilde{n}^*} \cdot \tilde{P}^{\tilde{n}^*} \cdot (1 - \tilde{P}\,)^{n - \tilde{n}^*} \cdot V(\tilde{n}^*\,) \cdot J(\tilde{D}_c\,) d\tilde{D}_c =$$

(12)

$$=(n+1)\sum_{\tilde{n}^*=0}^{n} C_n^{\tilde{n}^*}\cdot \tilde{P}^{\tilde{n}^*}(1-P)^{n-\tilde{n}^*}\cdot \int_{-\infty}^{\infty} V(\tilde{n}^*)\cdot J(\tilde{D}_c)d\tilde{D}_c\,.$$

Користуючись отриманим виразом (12), можна у кожному конкретному випадку вимірювань визначити функцію достовірності інтервальної оцінки $P[X\in(x',x'')]$ випадкової величини X [8–10]. Цей вираз дає можливість оцінити функцію розподілу $f(x)$ ймовірності X, для чого вісь x розбивається на рівні інтервали D_x і знаходиться функція достовірності для кожного інтервалу за формулою (12); при цьому утворюється функція двох змінних.

Висновки

Спосіб практичного використання функції достовірності $J(\tilde{P})$ залежить від характеру, завдань і критеріїв ефективності цієї системи [6–10].

Аналізуючи системи і використовуючи розрахункові співвідношення, до яких входять значення ймовірності, слід мати на увазі, що ці значення можуть бути отримані з результатів спостережень тільки як оцінкові. У розрахунках потрібно користуватися функціями достовірності оцінок, подібно до того, як в розрахунках, що містять випадкові величини, належить користуватися функціями розподілу ймовірності для отримання точнішого результату.

З введенням поняття функції достовірності випадкових процесів з'являється можливість об'єктивного кількісного оцінювання точності відновлення моделлю процесу.

1. Балашов В.П. Криксунов В.С., Сретенский В.Н. Системный анализ и основные принципы отраслевой метрологии // Электронная техника. Серия № 8. "Управление качеством и стандартизация", 1975. – Вып. 7(17). 2. Криксунов В.С. Об оценивании физических величин по результатам измерений / Труды войсковой части 55215, 2000. – Вып. 1(20). 3. ГОСТ 16265-70. 4. Прохоров Ю.В., Розанов Ю.А. Теория вероятностей. – М.: Наука, 1973. 5. Вентцель Е.С. Теория вероятностей. – М.: Наука, 1992. 6. Левченко А.О. Кількісні характеристики точності алгоритмів прогнозу // НТЗ. 2001. – №6. – С.46–52. 7. Левченко А.О. Засоби атестації систем метрологічного супроводження / Тези доповідей наук.-техн. конф. "Стан і розвиток військово-морських сил Збройних сил України на сучасному етапі", 2003. 8. Левченко А.О., Стадник И.Л. Анализ точности восстановления плотностей распределения вероятностей решением уравнения свертки методом трапеций // Тезисы докладов: Материалы XII Семінара "Моделирование в прикладных научных исследованиях".– Одесса: ОНПУ, 2005. 9. Левченко А.А., Стадник И.Л. Методы оценивания достоверности идентификации состояния информационных измерительных систем // Труды Одесского политехнического університета, 2006. – Вып. 1(20). 10. Левченко А.О., Кравчук О.І., Шаріпова І.В. Визначення помилки апроксимації моделі процесу зміни технічного стану технічних об'єктів на великих строках експлуатації // Тезисы докладов 6-й международной научно-практической конференции "Современные информационные и электронные технологии". – 2006.