УДК 621.331

В.Г. Кузнєцов, канд. техн. наук
(Україна, Дніпропетровськ, Дніпропетровський національний університет залізничного транспорту ім. акад. В. Лазаряна)

ВИЗНАЧЕННЯ ПОХИБОК І ТРИВАЛОСТІ СТАТИСТИЧНИХ ВИМІРІВ НАВАНТАЖЕНЬ ЕЛЕКТРОТЯГИ

Ймовірносно-статистичний підхід до вивчення навантажень електротяги, особливо їх максимальних значень, потребує проведення тривалих статистичних випробувань. Для цього використовують сучасні цифрові засоби виміру параметрів електричних навантажень.

Необхідна статистична інформація може бути отримана за допомогою дискретних вимірів. Звіт про них виводиться періодично з інтервалом Dt [2, 5].

Для такого роду приладів виникає необхідність розробки питань стосовно точності і надійності оцінок, отриманих при статистичних випробуваннях, тривалості експерименту і параметрах аналізуючого пристрою.

Основною умовою, що забезпечує застосування законів математичної статистики, є випадковість відбору елементів вибіркової сукупності. При періодичній вибірці вона може бути досягнута лише спостереженням некорельованих значень. При статистичному аналізі навантажень електротяги, враховуючи близькість закону розподілу до нормального, некорелованість їх вибіркових значень буде приблизно означати їх незалежність.

При статистичних вимірюваннях мають місце помилки методу вимірювання або репрезентативності і помилки реєстрації. Помилки репрезентативності можуть бути випадковими і систематичними.

Випадкові помилки спричиняються багаточисельними факторами, дія яких різна при повторних вимірах, і залежать від тривалості експерименту. Систематичні помилки мають місце, коли порушуються умови випадковості відбору. Їх джерело – квантування за рівнем і дискретизація за часом.

У даній роботі розглядаються лише помилки репрезентативності. Помилками реєстрації нехтуватимемо, вважаючи їх за нескінченно малі.

Випадкові похибки

При вивченні ймовірнісної природи навантажень електротяги підлягають визначенню похибки оцінок одновимірних розподілів ймовірності і числових характеристик.

Виконавши певні передумови (вибірка досить великого обсягу повинна складатися з незалежних і однаково розподілених елементів), оцінки можна вважати нормальними і характеризувати двома параметрами – математичним очікуванням і середньоквадратичним відхиленням.

Як випадкова величина оцінка характеристики варіюється від вибірки до вибірки, і судити про неї можна лише з ймовірності того, що вона не вийде за межі довірчого інтервалу.

Оцінкою генеральної ймовірності служить вибіркова частота

$$R = \frac{n_i}{n}, \tag{1}$$

де $-n_i$ – число відліків лічильника; i – рівень; n – сума відліків лічильників всіх рівнів.

Математичним очікуванням оцінки є генеральна ймовірність

$$\frac{P_0}{M_{[p]}} = P_0,$$

а середньоквадратичне відхилення [1] є її помилкою, тобто

$$s_p = \sqrt{\frac{P_0 \cdot (1 - P_0)}{n}}. \tag{2}$$

Довірчий інтервал оцінки P виявиться рівним $[P - \Delta p, P + \Delta p]$, де $Dp - ts_p$ – гранична помилка, t – аргумент, що відповідає довірчій ймовірності.

Задаючись похибкою оцінки s_p, неважко визначити об'єм вибірки n_p

$$n_p = \frac{P_0 \cdot (1 - P_0)}{s_p^2} \tag{3}$$

Звуження довірчого інтервалу, тобто зменшення граничної помилки Dp, приведе при тій самій точності оцінки s_p до різкого зменшення рівня достовірності b. Іншими словами для звуження довірчого інтервалу слід рекомендувати зменшення середньоквадратичної помилки s_p, тобто збільшення об'єму вибіркових даних.

Отже, з достовірністю b можна стверджувати, що шукана ймовірність P_0 знаходитиметься в межах довірчого інтервалу

$$P \cdot (p - Dp \le P_0 < P + Dp) = b \tag{4}$$

Шляхом аналогічних міркувань визначаються граничні помилки і відповідний об'єм вибіркових елементів для решти необхідних характеристик, які зведені у таблицю.

Відмітимо деякі особливості у визначенні похибок окремих оцінок і об'ємів вибірок. При визначенні об'єму вибірки генеральна ймовірність апріорі не відома. В цьому випадку використовується властивість максимуму добутку ймовірності подій, що складають повну групу:

$$max[P_0 \cdot (1 - P_0)] = 0.25. \tag{5}$$

Одновимірні закони розподілу зручніше оцінювати по їх відносних середньоквадратичних помилках s . Так, для функції розподілу зазвичай оцінюють значення $F(I)$.

$$d_F = \frac{s_F}{F(I)} \tag{6}$$

Величини граничних помилок та необхідного об'єму вибірки

Характеристика генеральної сукупності	Оцінка за вибірковими даними	Середньоквадратична похибка	Необхідний об'єм вибірки
Ймовірність	$p = \frac{n_i}{n}$	$s_p = \sqrt{\frac{p_0 \cdot (1 - p_0)}{n}}$	$n_p = \frac{p_0 \cdot (1 - p_0)}{s_p^2}$
Функція розпод.	$F(I) = \frac{1}{n} \sum_{i=1}^{j} n_i$	$s_F = \sqrt{\frac{F(I) \cdot (1 - F(I))}{n}}$	$n = \frac{p_0 \cdot (1 - p_0)}{s_p^2}$
Щільність розпод.	$f(I) = \frac{1}{n \cdot DI} \sum_{i=1}^{j} n_i$	$s_f = \sqrt{\frac{f(I)}{n \cdot f \cdot DI}}$	$n_f = \frac{1}{s_f^2 \cdot P_{D\,Imin}}$
Математичне очікування	$\bar{I} = \frac{1}{n} \sum_{i=1}^{m} I_i \cdot n_i$	$s_{\bar{I}} = \frac{s_0}{\sqrt{n_{\bar{I}}}}$	$n_{\bar{I}} = \frac{s_0^2}{s_{\bar{I}}^2}$
Дисперсія	$s^2 = \frac{1}{n} \sum_{i=1}^{m} (I_i - \bar{I}) \cdot n_i$	$s_{s^2} = \sqrt{\frac{2}{n} \cdot s_0^2}$	$n_{s^2} = \frac{2 \cdot s_0^4}{s_{s^2}}$
Стандарт	$s = \sqrt{\frac{1}{n} \sum_{i=1}^{m} (I_i - \bar{I}) \cdot n_i}$	$s_s = \frac{s_0}{\sqrt{2 \cdot p}}$	$n_s = \frac{s_0^2}{2 \cdot s_s^2}$
Асиметрія	$A = \frac{1}{n \cdot s^3} \sum_{i=1}^{m} (I_i - \bar{I})^3 \cdot n_i$	$s_A = \sqrt{\frac{6}{n}}$	$n_A = \frac{6}{s_{A^2}}$
Ексцес	$E = \frac{1}{n \cdot s^4} \sum_{i=1}^{m} (I_i - \bar{I})^4_{n_i} \cdot 3$	$s_E = 2 \cdot \sqrt{\frac{6}{n}}$	$n_E = \frac{24}{s_{E^2}}$

Примітка: DI – ширина каналу дискретизації

m – число каналів дискримінації

$P_{D\,Imin}$ – ймовірність попадання випадкової величини в крайній інтервал групування.

При оцінці щільності ймовірності цікавляться похибками на краях розподілу. Знаючи ймовірність попадання випадкового навантаження в крайній проміжок DI, неважко підрахувати і щільність ймовірності для цього проміжку

$$f(I)=\frac{P_{D\,Imin}}{DI}. \tag{7}$$

При цьому об'єм вибірки nf виявляється незалежним від щільності ймовірності $f(I)$ і ширини каналу DI, а залежить тільки від відносної похибки d_f і ймовірності $P_{D\,Imin}$. Це полегшує визначення nf, наприклад, порівняно з [3]. Оцінка похибок математичного очікування, дисперсії і стандарту пов'язана з генеральною дисперсією процесу s_0^2, яка сама підлягає визначенню. Щоб вийти з цього ускладнення, можна скористатися значенням вибіркової дисперсії s^2, що отримано раніше приблизно в аналогічних умовах, або організувати невелике вибіркове спостереження. Слід зазначити, що при однаковій точності і надійності оцінок дисперсії і стандарту для дисперсії потрібний об'єм вибіркової сукупності в 4 рази більший.

Оцінки асиметрії і ексцесу, що показують відмінність розподілу випадкових навантажень від нормального, вимагають при однаковій точності і надійності значно більшого об'єму вибіркових совокупностей, ніж оцінки нормального розподілу – математичне очікування і дисперсія.

Тривалість експерименту Т

Основні характеристики експерименту – тривалість і період дискретизації. Вони і визначають близькість отриманих результатів до дійсних характеристик випадкового процесу. Час спостереження визначається як

$$T = Dt \cdot n_{max}, \tag{8}$$

де Dt – період дискретизації; n_{max} – максимальний об'єм вибіркової сукупності.

Для некорельованої вибірки період дискретизації має бути більше періоду кореляції: $Dt \geq t_k$. Період кореляції визначимо з виразу для кореляційної залежності навантаження фідера [4], тобто

$$K_I(t)=s^2 \cdot exp(-\frac{t}{T_k}), \tag{9}$$

де T_k – постійна часу загасання кореляційних зв'язків.

$$T_k = \frac{T}{2 \cdot N}, \text{годин}; \tag{10}$$

N – пропускна здатність ділянки.

Практично при $K_I(t) = 0.05 \cdot s^2$ сусідні елементи вибірки можна вважати за некорельовані. Задавшись пропускною здатністю $N = 60$ п/добу, отримаємо $T_k = 36$ хв. Зі збільшенням пропускної здатності T_k зменшується. Отже, можна узяти і менший Dt.

Систематичні похибки

Це похибки дискретних вимірювань. Похибка дискретизації за часом виникає разом із втратою частини інформації. Ця похибка апроксимації, тобто різниця між точними і відновленим значенням виміряної величини [4], залежить від кроку Dt і властивостей процесу.

Математичне очікування цієї похибки при стаціонарному процесі $I(t)$ $M[h_a] = 0$, а дисперсія

$$Д[h_a] = 2 \cdot s^2 \cdot (-r_{t-s}), \tag{11}$$

де r_{t-s} – нормований коефіцієнт кореляції між моментами вимірювання.

У нашому випадку

$$r_{t-s} = 0.05;$$
$$s_{ra}^2 = 2 \cdot 0.95 \cdot s^2 = 1.9 \cdot s^2;$$
$$s_{ha} = 1.38 \cdot s_0.$$

Середньоквадратичне відхилення похибки апроксимації виходить достатньо великим за рахунок некорелованості вибіркових значень та із зменшенням s процесу зменшуватиметься.

Похибка квантування по рівню зачіпає центральні моменти парних порядків, тобто дисперсію та ексцес, виникає від округлення вибіркових значень і вимагає віднесення їх до середини інтервалу групування. Математичне очікування її $M[h_a] = 0$ а середньоквадратичне відхилення $s_{hy} = \dfrac{DI}{2 \cdot \sqrt{3}}$.

Число каналів дискретизації

Число каналів дискретизації визначається виходячи із закону розподілу процесу[2]. Враховуючи близькість останнього до нормального, отримаємо:

$$m = \frac{1505}{4 \cdot \sqrt{d_y}} + 1, \tag{12}$$

де $d_y = \dfrac{Df(I)}{f(I)}$ - допустима відносна похибка апроксимації щільності ймовірності.

Приймемо $\delta_y = 0{,}1$. Тоді $m = \dfrac{1505}{4 \cdot \sqrt{0.1}} + 1 = 11$.

З погляду на зручність розбиття інтервалу вимірювання на 10 можна прийняти $m = 10$. Крок квантування $DI = \dfrac{1000}{10} = 100$ А.

Максимальна ємність лічильника

Максимальна ємність лічильника визначається з міркувань вибору найбільш завантаженого каналу з найбільшою щільністю ймовірності

$$f(I)_{max} = \frac{1}{s \cdot \sqrt{2 \cdot f_1}}. \tag{13}$$

При $s = 70$ отримаємо

$$f(I) = \frac{1}{70 \cdot \sqrt{6.28}} = 0.0056, \tag{14}$$

а ймовірність попадання в цей канал

$$P_{max} = f(I)_{max} \cdot DI; \tag{15}$$

$$P_{max} = 0.0056 \cdot 100 = 0.56. \tag{16}$$

Ємність найбільш завантаженого каналу

$$K = P_{max} \cdot n, \tag{17}$$

де $n = 10^4$ – максимальний об'єм вибірки, тоді

$$K = 0{,}56 \cdot 10^4 = 5600. \tag{18}$$

Отже, для вимірювань параметрів тягового навантаження достатньо 4-х розрядних лічильників.

Висновки

– Надано класифікацію похибок, що виникають при експериментальному визначенні параметрів системи тягового електропостачання.

– Наведено підхід до визначення похибок і тривалості статистичних вимірів навантажень електротяги з урахування ймовірнісної природи процесу електроспоживання тяги поїздів.

Список літератури

1. Андронов А.М., Копытов Е.А., Гринглаз Л.Я. Теория вероятностей и математическая статистика. С.Пб.: Узд. дом "Питер". – 2004. – 460 с.

2. Сиченко В.Г., Кузнецов В.Г., Босий Д.О. Дослідження режиму напруги в тяговій мережі змінного струму при рекуперації електроенергії // Вісн. Призов. держ. техн. ун-ту. – 2008. Вип. № 18. ч. 2. – С.35-39
3. Немировский А. С. Вероятностные методы в измерительной технике. М.: Изд-во стандартов. – 1969. – 86 с.
4. Г. Я. Мирский. Аппаратурное определение характеристик случайных процессов. М.: Энергия, 1967. – 114 с.
5. Прибор для измерений качества и учёта электрической энергии PM 175. Руководство по установке и эксплуатации. – М. – 155 с.