АКАДЕМИЯ НАУК СССР
СИБИРСКОЕ ОТДЕЛЕНИЕ
АВТОМЕТРИЯ

№ 3 1967

УДК 681.2.08

З. А. ЛИВШИЦ, В. И. РАБИНОВИЧ, О. Е. ТРОФИМОВ
(Новосибирск)

ОБ ИНФОРМАЦИОННОМ ПОДХОДЕ К ОПРЕДЕЛЕНИЮ ПАРАМЕТРОВ ЗАКОНА РАСПРЕДЕЛЕНИЯ ИЗМЕРЯЕМОЙ ВЕЛИЧИНЫ*

В течение ряда лет информационная оценка процессов и средств измерения привлекает внимание специалистов (см., например, [1—4]). Хотя указанные работы и отличаются подходом к определению количества измерительной информации и точками приложения информационного анализа, но имеют много общего. Одной из таких общих черт является предположение о том, что закон распределения вероятностей измеряемой величины полностью известен. Это характерно и для других методов исследования средств измерения. Именно этим обстоятельством объясняется появление в последние годы обширной литературы, посвященной вопросам проектирования приборов для определения законов распределения и их параметров.

Очевидно, что уровень априорных знаний о законе распределения измеряемой величины может быть различным. В одних случаях известны как вид закона распределения (нормальный, равномерный и т. д.), так и значения всех определяющих его параметров. В других случаях известны значения некоторых моментов (математического ожидания, дисперсии), но вид закона не определен. В настоящей статье анализируется ситуация, характерная тем, что вид закона распределения задан, но значения одного или нескольких его параметров неизвестны.

Обычным аппаратом в таких случаях являются методы математической статистики. В [5] (по-видимому, впервые) рассмотренная ситуация изучается в рамках теории информации. Однако автор исходит из предположения о том, что априори известно распределение вероятностей возможных значений неизвестного параметра, что, видимо, не имеет места в большинстве практически важных задач.

Сформулируем теперь подробнее рассматриваемую задачу. Множество результатов измерения является выборочной совокупностью, на основании которой вычисляется значение оценки для неизвестного параметра. (Например, в качестве оценки неизвестного математического ожидания измеряемой величины используется среднее арифметическое выборочных значений.) Известно, что выборочные характеристики (в частности, оценки, основанные на выборках) являются случайными

---

* Материал доложен на VIII Всесоюзной конференции по автоматическому контролю и методам электрических измерений в сентябре 1966 года в Новосибирске.

величинами, распределение которых полностью определяется законом распределения «генеральной совокупности» — измеряемой величины. Поэтому разумность выбора реализовавшегося значения выборочной характеристики в качестве оценки параметра может количественно характеризоваться неопределенностью — энтропией указанной случайной величины. В самом деле, малая неопределенность выборочной характеристики означает, что с большой вероятностью в качестве оценки мы воспользуемся величиной, близкой к истинному значению параметра.

В дальнейшем мы будем исходить из следующих предположений:

1) прибор является «идеальным», т. е. отсутствует случайная погрешность;

2) в качестве меры неопределенности используется дифференциальная энтропия [6], т. е. квантование не учитывается.

Эти ограничения физически эквивалентны предположению о пренебрежимой малости погрешности измерения и необходимы единственно для того, чтобы вычисления, которые несколько усложнились бы без учета этих ограничений, не затемняли бы качественную сторону вопроса.

Изучим сначала случай, когда измеряемая величина x распределена нормально. Введем естественное предположение о том, что вероятность реализации значений случайной величины вне диапазона измерения близка к нулю.

Итак, пусть величина x распределена нормально со стандартным отклонением σ и неизвестным математическим ожиданием m. Введем следующие обозначения: $I_n(x)$ — среднее значение количества информации об измеряемой величине, получаемое в результате измерения; $H(p_n)$ — энтропия выборочной характеристики p_n, основанной на выборке из n значений измеряемой величины. В нашем случае [6]

$$I_n(x) = \log \sqrt{2\pi e}\,\sigma. \tag{1}$$

Из (1) следует, что дифференциальная энтропия случайной величины, распределенной по нормальному закону, не зависит от математического ожидания.

Для вычисления энтропии оценки заметим, что величина $\bar{x}_n = \frac{1}{n}\sum_{i=1}^{n} x_i$ — несмещенная оценка математического ожидания — распределена нормально с параметрами $\left(m, \frac{\sigma}{\sqrt{n}}\right)$ (см., например, [7]). Здесь x_i — результат i-го измерения; n — количество измерений (объем выборки).

Следовательно,

$$H(\bar{x}_n) = \log \sqrt{2\pi e}\,\frac{\sigma}{\sqrt{n}}. \tag{2}$$

Формула (2) показывает, что с ростом n энтропия оценки для математического ожидания уменьшается. Выясним, какой «вклад» вносит в уменьшение неопределенности каждое измерение. Обозначая через i изменение энтропии оценки в результате проведения n-го измерения, получим

$$i_n = H(\bar{x}_{n-1}) - H(\bar{x}_n) = \log \frac{\sqrt{n}}{\sqrt{n-1}}. \tag{3}$$

При больших n $i_h \sim \frac{1}{2(n-1)}$, т. е. при большом объеме выборки, выполнение еще одного измерения практически не приносит информации.

Рассмотрим теперь случай, когда математическое ожидание нормально распределенной измеряемой величины известно, но неизвестна дисперсия σ^2. В [7] показано, что несмещенной оценкой дисперсии является величина $y_n = \frac{n}{n-1} s_n^2$, где $s_n^2 = \frac{1}{n} \sum_{i=1}^{n} (\bar{x}_n - x_i)^2$ — выборочная дисперсия; y_n является случайной величиной, имеющей плотность распределения вероятностей

$$p(y_n) = \frac{n-1}{\sigma^2} K_{n-1}\left(\frac{(n-1)\,y_n}{\sigma^2}\right). \tag{4}$$

Здесь

$$K_{n-1}(z) = \frac{1}{2^{\frac{n-1}{2}}\,\Gamma\left(\frac{n-1}{2}\right)}\, z^{\frac{n-1}{2}-1}\, e^{-\frac{z}{2}}. \tag{5}$$

Таким образом,

$$p(y_n) = \frac{(n-1)^{\frac{n-1}{2}}}{\sigma^{n-1}\, 2^{\frac{n-1}{2}}\,\Gamma\left(\frac{n-1}{2}\right)}\, y_n^{n-3}\, e^{-\frac{(n-1)\,y_n}{2\sigma^2}}. \tag{6}$$

Логарифмируя (6), получаем

$$\ln p(y_n) = \frac{n-1}{2}\ln(n-1) - (n-1)\ln\sigma - \frac{n-1}{2}\ln 2 - $$
$$- \ln\Gamma\left(\frac{n-1}{2}\right) + \frac{n-3}{2}\ln y_n - \frac{n-1}{2\sigma^2}\, y_n. \tag{7}$$

Воспользуемся одним из представлений логарифма гамма-функции

$$\ln\Gamma(t) = \left(t - \frac{1}{2}\right)\ln t - t + \frac{1}{2}\ln 2\pi + R(t), \tag{8}$$

где

$$R(t) = \frac{1}{12\,t} + 0\left(\frac{1}{t^3}\right).$$

Из (7) и (8) находим

$$\ln p(y_n) = \frac{1}{2}\ln(n-1) - \frac{1}{2}\ln 2 - (n-1)\ln\sigma +$$
$$+ \frac{n-1}{2} - \frac{1}{2}\ln 2\pi - \frac{n-3}{2}\ln y_n + \frac{n-1}{2\sigma^2}\, y_n + R\left(\frac{n-1}{2}\right). \tag{9}$$

Вычислим теперь энтропию распределения случайной величины y_n:

$$H(y_n) = -\int_0^\infty p(y_n)\ln p(y_n)\,dy_n =$$

$$= -\frac{1}{2}\ln(n-1) + \frac{1}{2}\ln 2 + (n-1)\ln\sigma - \frac{n-1}{2} +$$

$$+ \frac{1}{2}\ln 2\pi - \frac{n-3}{2}\int_0^\infty p(y_n)\ln y_n\,dy_n + \frac{n-1}{2\sigma^2}\times$$

$$\times\int_0^\infty y_n p(y_n)\,dy_n + R\left(\frac{n-1}{2}\right). \qquad (10)$$

Выбор пределов интегрирования определяется тем, что величина $y_n = \frac{n}{n-1}s_n^2$ неотрицательна. Вычислим оставшиеся интегралы:

$$I_1 = \frac{n-1}{2\sigma^2}\int_0^\infty y_n p(y_n)\,dy_n = \frac{n-1}{2\sigma^2}M(y_n) = \frac{n-1}{2}, \qquad (11)$$

так как $M(y_n)$ есть σ^2;

$$I_2 = \frac{n-3}{2}\int_0^\infty p(y_n)\ln y_n\,dy_n = \frac{n-3}{2}\,\frac{(n-1)^{\frac{n-1}{2}}}{\sigma^{n-1}2^{\frac{n-1}{2}}\Gamma\left(\frac{n-1}{2}\right)}\times$$

$$\times\int_0^\infty y_n^{n-3}e^{\frac{1-n}{2\sigma^2}y_n}\ln y_n\,dy_n. \qquad (11a)$$

Воспользуемся соотношением [8]

$$\int_0^\infty x^{\nu-1}e^{-\mu x}\ln x\,dx = \frac{1}{\mu^\nu}\Gamma(\nu)\left[\psi(\nu) - \ln\mu\right], \qquad (12)$$

в котором $\psi(\nu) = \frac{d}{d\nu}\ln\Gamma(\nu)$.

В нашем случае $\nu = \frac{n-1}{2}$, а $\mu = \frac{n-1}{2\sigma^2}$. Дифференцируя (8) (дифференцирование правомерно ввиду равномерной сходимости $R(t)$ и $R'(t)$ на промежутке $[\varepsilon, \infty]$, $(0<\varepsilon)$), получим

$$\psi(\nu) = \ln\nu - \frac{1}{2\nu} + o\left(\frac{1}{\nu^2}\right);$$

$$\psi\left(\frac{n-1}{2}\right) = \ln\frac{n-1}{2} - \frac{1}{n-1} + o\left(\frac{1}{n^2}\right). \qquad (13)$$

Подстановка (12) и (13) в (11a) дает

$$I_2 = \frac{n-3}{2}\,\frac{n-1^{\frac{n-1}{2}}}{\sigma^{n-1}2^{\frac{n-1}{2}}\Gamma\left(\frac{n-1}{2}\right)}\,\frac{\Gamma\left(\frac{n-1}{2}\right)}{\left(\frac{n-1}{2\sigma^2}\right)^{\frac{n-1}{2}}}\times$$

$$\times\left[\ln(n-1)-\frac{1}{n-1}-\ln\frac{n-1}{2}+2\ln\sigma+o\left(\frac{1}{n}\right)\right]=$$

$$=(n-3)\ln\sigma-\frac{n-3}{2(n-1)}+o\left(\frac{1}{n}\right). \qquad (14)$$

Подставляя значения вычисленных интегралов в (10), получаем

$$H(y_n)=-\frac{1}{2}\ln(n-1)+2\ln\sigma+\frac{1}{2}\ln 2+\frac{1}{2}\ln 2\pi-$$

$$-\frac{n-3}{2(n-1)}+o\left(\frac{1}{n}\right)=\ln\sqrt{2\pi}\,\frac{\sqrt{2}}{\sqrt{n-1}}\sigma^2+\frac{1}{2}\,\frac{n-3}{n-1}+o\left(\frac{1}{n}\right). \qquad (15)$$

В формуле (15) фигурирует априори неизвестная величина σ. Однако, учитывая, что $\frac{n}{n-1}s_n^2$ сходится по вероятности к σ^2, и подставляя в (15) вместо σ^2 эмпирическое значение y_n, при достаточно больших n получим с вероятностью, близкой к единице, значения энтропии, близкие к истинным, т. е.

$$H(y_n)\approx\ln 2\sqrt{\pi(n-1)}\,\frac{n}{n-1}\,s_n^2+\frac{1}{2}\,\frac{n-3}{n-1}. \qquad (16)$$

Легко видеть, что величина $i_n(y_n)$, как и в первом рассматриваемом случае, имеет порядок $\frac{1}{n}$. В самом деле,

$$i_n(y_n)=H(y_{n-1})-H(y_n)\approx\ln\sqrt{\frac{n-2}{n-1}}\,\frac{n}{n-1}-$$

$$-\frac{1}{(n-1)(n-2)}\simeq\frac{1}{n}. \qquad (17)$$

Наконец, рассмотрим случай, когда неизвестны оба параметра распределения измеряемой величины — математическое ожидание и дисперсия. Соотношения, полученные для второго из изученных случаев, будут справедливы и в данной ситуации, так как энтропия случайных величин $\bar{x}_n$ и y не зависит от значения математического ожидания. Поскольку случайные величины $\bar{x}_n$ и y_n независимы [7], энтропия системы этих двух величин равна сумме их энтропий.

Столь подробное исследование случая, когда измеряемая величина распределена нормально, неслучайно. Дело в том, что для других законов точное распределение выборочных характеристик (за исключением среднего) почти не изучено. В таких ситуациях приходится пользоваться асимптотическими распределениями оценок. В [9] показано, что при весьма общих условиях возможен предельный переход под знаком энтропии. Это означает, что при достаточно больших n энтропия асимптотического распределения является хорошим приближением к истинной. Проиллюстрируем этот факт примером. Величина y_n, введенная выше, распределена асимптотически нормально с параметрами

$\left(\sigma^2, \frac{\sqrt{2n}\,\sigma^2}{n-1}\right)$. Энтропия асимптотического распределения вычисляется по формуле

$$H_{n\,\text{асимпт}} = \ln \sqrt{2\pi e\,\frac{\sqrt{2n}}{n-1}\,\sigma^2}. \tag{18}$$

Сравнивая (18) с (15), видим, что

$$\lim_{n \to \infty} \left(H(y_n) - H_{n\,\text{асимпт}}\right) = 0. \tag{19}$$

ЛИТЕРАТУРА

1. П. В. Новицкий. Использование кибернетических понятий в теории электроизмерительных устройств.— Измерительная техника, 1962, № 1.
2. В. И. Рабинович, М. П. Цапенко. О количестве измерительной информации.— Измерительная техника, 1963, № 4.
3. М. И. Ланин, С. М. Мандельштам, В. В. Сидельников. Некоторые вопросы математического обоснования выбора числа областей квантования в аналого-дискретных преобразователях.— Автоматика и телемеханика, 1963, № 4.
4. С. М. Персин. Количество информации при цифровом измерении.— Измерительная техника, 1964, № 7.
5. Д. В. Линдли. О мере информации, даваемой экспериментом. Математика, 1959, № 3 : 3.
6. К. Шеннон. Работы по теории информации и кибернетике. М.—Л., Изд-во иностр. лит., 1963.
7. Г. Крамер. Математические методы статистики. М.—Л., Изд-во иностр. лит., 1948.
8. И. С. Градштейн, И. М. Рыжик. Таблицы интегралов, сумм, рядов и произведений. Изд. 4-е. М., Физматгиз, 1962.
9. Р. Л. Добрушин. Предельный переход под знаком энтропии.— Теория вероятностей и ее применения, 1960, т. V, вып. 1.

Поступила в редакцию
6 августа 1966 г.,
окончательный вариант —
21 ноября 1966 г.