УДК 519.22:681.2.08

В.С. Берегун, канд. техн. наук, **О.В. Гармаш,**
Национальный технический университет Украины «КПИ»
(Украина, 03056 Киев, ул. Политехническая, 16, корп. 12,
тел. (044) 4549072, e-mail: viktorberegun@i.ua; oks.garmash@gmail.com),
А.И. Красильников, канд. физ.-мат. наук
Ин-т технической теплофизики НАН Украины,
(Украина, 03057 Киев, ул. Желябова, 2а,
тел. (044) 4532857, e-mail: tangorov@voliacable.com)

Среднеквадратические ошибки оценок кумулянтных коэффициентов пятого и шестого порядков

Получены общие выражения для определения математического ожидания и дисперсии оценок кумулянтных коэффициентов пятого и шестого порядков, которые позволяют вычислять среднеквадратические ошибки оценок этих коэффициентов. Проанализированы ошибки оценок кумулянтных коэффициентов γ_s, $s = \overline{3,6}$, некоторых типовых симметричных распределений при фиксированном объеме выборки. Определен минимальный объем выборки, при котором обеспечивается наперед заданная относительная ошибка оценок кумулянтных коэффициентов этих распределений.

Отримано загальні вирази для визначення математичного сподівання та дисперсії оцінок кумулянтних коефіцієнтів п'ятого та шостого порядків, які дозволяють обчислювати середньоквадратичні помилки оцінок цих коефіцієнтів. Проаналізовано помилки оцінок кумулянтних коефіцієнтів γ_s, $s = \overline{3,6}$, деяких типових симетричних розподілів при фіксованому об'ємі вибірки. Визначено мінімальний об'єм вибірки, при якому забезпечується наперед задана відносна помилка оцінок кумулянтних коефіцієнтів цих розподілів.

Ключевые слова: оценка, моменты, кумулянтные коэффициенты, среднеквадратическая ошибка, относительная ошибка.

В настоящее время решение многих прикладных задач [1—3], основанных на вероятностном подходе, базируется на негауссовских моделях исследуемых случайных величин и случайных процессов. При анализе негауссовских распределений большое значение имеют кумулянтные коэффициенты $\gamma_s = \kappa_s / \kappa_2^{s/2}$. Здесь κ_s — кумулянты распределения [4, 5],

$$\kappa_s = \left. \frac{d^s \ln f(u)}{i^s du^s} \right|_{u=0},$$

где $f(u)$ — характеристическая функция; $i = \sqrt{-1}$.

Из всех кумулянтных коэффициентов чаще всего используются коэффициенты асимметрии γ_3 и эксцесса γ_4, являющиеся простыми и удобными числовыми характеристиками, показывающими степень отличия исследуемого распределения от гауссовского, для которого $\gamma_s = 0$ при $s \geq 3$. С использованием коэффициентов γ_3 и γ_4 осуществляется выбор аппроксимирующих плотностей вероятности из систем распределений Пирсона [6] и Джонсона [7]. В работе [8] коэффициенты γ_3 и γ_4 использованы в качестве информативных параметров для диагностики подшипников качения. Однако использование при решении задач только коэффициентов асимметрии γ_3 и эксцесса γ_4 может привести к ошибочным результатам. В частности, нетрудно убедится, что значениям $\gamma_3 = 0$ и $\gamma_4 = 3$ соответствуют распределения Лапласа и Стьюдента с параметром $\nu = 6$, свойства которых существенно различны.

Кроме того, при решении ряда практических задач возникает необходимость использования кумулянтных коэффициентов выше четвертого порядка. Например, коэффициенты γ_s, $s = \overline{3,6}$, необходимы для аппроксимации негауссовской плотности вероятности отрезками ортогональных рядов [9], гауссовской смесью распределений [10], приближенного нахождения пуассоновской спектральной плотности безгранично делимых случайных величин [11]. В работе [12] исследована возможность применения кумулянтных коэффициентов γ_s, $s = \overline{3,6}$, для ранней диагностики заболеваний коленных суставов.

При практическом применении кумулянтных коэффициентов используются их оценки $\hat{\gamma}_s$, получаемые по экспериментальным данным. Поскольку любые оценки являются случайными величинами, необходимо знать ошибки оценок, в частности среднеквадратические ошибки, выражаемые через математическое ожидание и дисперсию оценок. В работах [4, 5] получены математическое ожидание и дисперсия оценок кумулянтных коэффициентов γ_s, $s = 3, 4$. Для $s \geq 5$ такие результаты неизвестны.

Определим математическое ожидание и дисперсии оценок кумулянтных коэффициентов γ_s, $s = 5, 6$, что позволит находить среднеквадратические ошибки оценок этих коэффициентов.

Постановка задачи и основные исходные формулы. Пусть имеется случайная величина ξ, у которой существуют кумулянты κ_s, $s = \overline{1,12}$. В результате проведения эксперимента со случайной величиной ξ получаем выборку $\xi_1, ..., \xi_N$ объема N, элементы которой являются независимыми одинаково распределенными случайными величинами. Необходимо получить формулы для оценок кумулянтных коэффициентов γ_s, $s = 5, 6$, и найти среднеквадратические ошибки этих оценок.

Для получения оценок параметров случайной величины используются различные методы [5, 13, 14], в частности метод моментов, метод максимального правдоподобия, метод наименьших квадратов, метод максимизации полинома и др. Формулы для оценок кумулянтных коэффициентов γ_s, $s=5,6$, несложно получить по методу моментов, используя известные соотношения [4, 5], выражающие кумулянтные коэффициенты через центральные моменты:

$$\hat{\gamma}_5 = \frac{\hat{\mu}_5 - 10\hat{\mu}_3\hat{\mu}_2}{(\hat{\mu}_2)^{5/2}}, \tag{1}$$

$$\hat{\gamma}_6 = \frac{\hat{\mu}_6 - 15\hat{\mu}_4\hat{\mu}_2 - 10(\hat{\mu}_3)^2}{(\hat{\mu}_2)^3} + 30, \tag{2}$$

где $\hat{\mu}_s$ — оценки центральных моментов,

$$\hat{\mu}_s = \frac{1}{N}\sum_{k=1}^{N}(\xi_k - \hat{m})^s, \quad \hat{m} = \frac{1}{N}\sum_{k=1}^{N}\xi_k .$$

В математической статистике [4, 5, 13, 15] основными характеристиками качества оценок любого параметра θ случайной величины ξ являются несмещенность и состоятельность оценки $\hat{\theta}$. Оценка $\hat{\theta}$ параметра θ называется несмещенной, если ее математическое ожидание равно θ при любом объеме выборки, т.е. $\mathbf{M}\{\hat{\theta}\} = \theta$. Если данное условие не выполняется, то оценка $\hat{\theta}$ является смещенной, а величина $\Delta\{\hat{\theta}\} = \mathbf{M}\{\hat{\theta}\} - \theta$ называется смещением оценки. Если при увеличении объема выборки смещение стремится к нулю, то оценка является асимптотически несмещенной.

Оценка $\hat{\theta}$ называется состоятельной, если при увеличении объема выборки она сходится по вероятности к параметру θ. Критерием состоятельности является выполнение условия $\lim\limits_{N\to\infty} \mathbf{D}\{\hat{\theta}\} = 0$. В технических задачах [15, 16] характеристикой точности оценки $\hat{\theta}$ параметра θ является среднеквадратическая ошибка

$$\varepsilon = [\mathbf{D}\{\hat{\theta}\} + \Delta^2\{\hat{\theta}\}]^{1/2}, \tag{3}$$

в которой дисперсия $\mathbf{D}\{\hat{\theta}\}$ характеризует случайную ошибку оценки, а смещение $\Delta\{\hat{\theta}\}$ — ее систематическую ошибку. При условии $\theta \neq 0$ удобно рассматривать относительную ошибку оценки $\hat{\theta}$ в виде

$$\delta = \varepsilon / \theta. \tag{4}$$

Для нахождения математического ожидания и дисперсии оценок (1) и (2), входящих в формулу (3), воспользуемся известной теоремой [4], суть которой заключается в следующем. Пусть $H(\hat{\mu}_1, ..., \hat{\mu}_k)$ — функция от оценок моментов. Тогда

$$\mathbf{M}\{H(\hat{\mu}_1, ..., \hat{\mu}_k)\} = H(\mu_1, ..., \mu_k) + O(1/N), \tag{5}$$

$$\mathbf{D}\{H(\hat{\mu}_1, ..., \hat{\mu}_k)\} = \sum_{j,l=1}^{k} H_j H_l \, \mu_{11}\{\hat{\mu}_l, \hat{\mu}_j\} + O(1/N^{3/2}), \tag{6}$$

где H_j — значения первых частных производных функции $H(\hat{\mu}_1, ..., \hat{\mu}_k)$ в точках $\hat{\mu}_j = \mu_j$, $j = \overline{1, k}$; $\mu_{11}\{\hat{\mu}_l, \hat{\mu}_j\}$ — смешанный второй центральный момент оценок центральных моментов l-го и j-го порядков,

$$\mu_{11}\{\hat{\mu}_l, \hat{\mu}_j\} = \frac{\mu_{l+j} - l\mu_{l-1}\mu_{j+1} - j\mu_{l+1}\mu_{j-1} - \mu_l\mu_j + lj\mu_2\mu_{l-1}\mu_{j-1}}{N} + O\left(\frac{1}{N^2}\right); \tag{7}$$

$O(g(x))$ — величина того же порядка, что и функция $g(x)$. В случае, когда порядки равны, $l = j$, формула (7) примет вид

$$\mu_{11}\{\hat{\mu}_l, \hat{\mu}_l\} = \mu_2\{\hat{\mu}_l\} = \frac{\mu_{2l} - 2l\mu_{l-1}\mu_{l+1} - \mu_l^2 + l^2\mu_2\mu_{l-1}^2}{N} + O\left(\frac{1}{N^2}\right), \tag{8}$$

где $\mu_2\{\hat{\mu}_l\}$ — дисперсия оценки центрального момента l-го порядка.

Учитывая, что при оценивании моментов требуются большие объемы выборок N, в дальнейшем вторыми слагаемыми в формулах (5)—(8) будем пренебрегать.

Ошибки оценки кумулянтного коэффициента $\boldsymbol{\gamma_5}$. Математическое ожидание оценки (1) коэффициента γ_5 на основании (5) имеет вид

$$\mathbf{M}\{\hat{\gamma}_5\} = \frac{\mu_5 - 10\mu_3\mu_2}{(\mu_2)^{5/2}} = \gamma_5, \tag{9}$$

откуда следует, что оценка (1) является несмещенной, т.е. систематическая ошибка отсутствует.

Для нахождения случайной ошибки оценки (1) определим ее дисперсию. Запишем выражение (6) с учетом (1):

$$\mathbf{D}\{\hat{\gamma}_5\} = H_1^2\mu_2\{\hat{\mu}_2\} + H_2^2\mu_2\{\hat{\mu}_3\} + H_3^2\mu_2\{\hat{\mu}_5\} + 2H_1H_2\mu_{11}\{\hat{\mu}_2, \hat{\mu}_3\} +$$
$$+ 2H_1H_3\mu_{11}\{\hat{\mu}_2, \hat{\mu}_5\} + 2H_2H_3\mu_{11}\{\hat{\mu}_3, \hat{\mu}_5\}. \tag{10}$$

Найдем составляющие выражения (10), используя формулы (7) и (8):

$$H_1 = \left.\frac{\partial H(\hat{\mu}_2, \hat{\mu}_3, \hat{\mu}_5)}{\partial \hat{\mu}_2}\right|_{\hat{\mu}_s = \mu_s} = \frac{30\mu_2\mu_3 - 5\mu_5}{2\mu_2^{7/2}},$$

$$H_2 = \frac{\partial H(\hat{\mu}_2, \hat{\mu}_3, \hat{\mu}_5)}{\partial \hat{\mu}_3}\bigg|_{\hat{\mu}_s = \mu_s} = -\frac{10}{\mu_2^{3/2}}, \; H_3 = \frac{\partial H(\hat{\mu}_2, \hat{\mu}_3, \hat{\mu}_5)}{\partial \hat{\mu}_5}\bigg|_{\hat{\mu}_s = \mu_s} = \frac{1}{\mu_2^{5/2}}; \quad (11)$$

$$\mu_2\{\hat{\mu}_2\} = \frac{\mu_4 - \mu_2^2}{N}, \quad (12)$$

$$\mu_2\{\hat{\mu}_3\} = \frac{\mu_6 - 6\mu_2\mu_4 - \mu_3^2 + 9\mu_2^3}{N}, \quad (13)$$

$$\mu_2\{\hat{\mu}_5\} = \frac{\mu_{10} - 10\mu_4\mu_6 - \mu_5^2 + 25\mu_2\mu_4^2}{N}, \quad (14)$$

$$\mu_{11}\{\hat{\mu}_2, \hat{\mu}_3\} = \frac{\mu_5 - 4\mu_2\mu_3}{N}, \quad (15)$$

$$\mu_{11}\{\hat{\mu}_2, \hat{\mu}_5\} = \frac{\mu_7 - 5\mu_3\mu_4 - \mu_2\mu_5}{N}, \quad (16)$$

$$\mu_{11}\{\hat{\mu}_3, \hat{\mu}_5\} = \frac{\mu_8 - 5\mu_4^2 - 3\mu_2\mu_6 - \mu_3\mu_5 + 15\mu_2^2\mu_4}{N}. \quad (17)$$

Обозначим через M_s нормированные центральные моменты:

$$M_s = \mu_s / \mu_2^{s/2}. \quad (18)$$

Подставляя выражения (11)—(17) в формулу (10) и учитывая (18), после преобразований получаем

$$\mathbf{D}\{\hat{\gamma}_5\} = c_5 / N, \quad (19)$$

где

$$c_5 = 75M_3^2M_4 + 875M_3^2 - 50M_3M_4M_5 - 435M_3M_5 +$$
$$+ \frac{25}{4}M_4M_5^2 + \frac{191}{4}M_5^2 + 900 + 160M_6 - 900M_4 +$$
$$+ M_{10} - 10M_4M_6 + 125M_4^2 + 30M_3M_7 - 5M_5M_7 - 20M_8. \quad (20)$$

Из выражения (19) следует, что оценка (1) является состоятельной.

Для симметричных распределений выражения (9) и (20) принимают следующий вид:

$$\mathbf{M}\{\hat{\gamma}_5\} = 0,$$

$$c_5 = 160M_6 - 900M_4 + 900 + M_{10} - 10M_4M_6 + 125M_4^2 - 20M_8. \quad (21)$$

Ошибки оценки кумулянтного коэффициента $\boldsymbol{\gamma_6}$. Математическое ожидание оценки (2) коэффициента γ_6 на основании (5) запишем в виде

$$\mathbf{M}\{\hat{\gamma}_6\}=\frac{\mu_6-15\mu_4\mu_2-10(\mu_3)^2}{(\mu_2)^3}+30=\gamma_6, \tag{22}$$

откуда следует, что оценка (2) является несмещенной, т.е. систематическая ошибка отсутствует.

Для нахождения случайной ошибки оценки (2) получим ее дисперсию. Запишем выражение (6) с учетом формулы (2):

$$\mathbf{D}\{\hat{\gamma}_6\}=H_1^2\mu_2\{\hat{\mu}_2\}+H_2^2\mu_2\{\hat{\mu}_3\}+H_3^2\mu_2\{\hat{\mu}_4\}+H_4^2\mu_2\{\hat{\mu}_6\}+$$
$$+2H_1H_2\mu_{11}\{\hat{\mu}_2,\hat{\mu}_3\}+2H_1H_3\mu_{11}\{\hat{\mu}_2,\hat{\mu}_4\}+2H_1H_4\mu_{11}\{\hat{\mu}_2,\hat{\mu}_6\}+$$
$$+2H_2H_3\mu_{11}\{\hat{\mu}_3,\hat{\mu}_4\}+2H_2H_4\mu_{11}\{\hat{\mu}_3,\hat{\mu}_6\}+2H_3H_4\mu_{11}\{\hat{\mu}_4,\hat{\mu}_6\}. \tag{23}$$

Найдем составляющие выражения (23), используя формулы (7) и (8):

$$H_1=\left.\frac{\partial H(\hat{\mu}_2,\hat{\mu}_3,\hat{\mu}_4,\hat{\mu}_6)}{\partial\hat{\mu}_2}\right|_{\hat{\mu}_s=\mu_s}=\frac{-3\mu_6+30\mu_2\mu_4+30\mu_3^2}{\mu_2^4},$$

$$H_2=\left.\frac{\partial H(\hat{\mu}_2,\hat{\mu}_3,\hat{\mu}_4,\hat{\mu}_6)}{\partial\hat{\mu}_3}\right|_{\hat{\mu}_s=\mu_s}=-\frac{20\mu_3}{\mu_2^3},$$

$$H_3=\left.\frac{\partial H(\hat{\mu}_2,\hat{\mu}_3,\hat{\mu}_4,\hat{\mu}_6)}{\partial\hat{\mu}_4}\right|_{\hat{\mu}_s=\mu_s}=-\frac{15}{\mu_2^2}, \tag{24}$$

$$H_4=\left.\frac{\partial H(\hat{\mu}_2,\hat{\mu}_3,\hat{\mu}_4,\hat{\mu}_6)}{\partial\hat{\mu}_6}\right|_{\hat{\mu}_s=\mu_s}=\frac{1}{\mu_2^3};$$

$$\mu_2\{\hat{\mu}_4\}=\frac{\mu_8-8\mu_3\mu_5-\mu_4^2+16\mu_2\mu_3^2}{N},$$

$$\mu_2\{\hat{\mu}_6\}=\frac{\mu_{12}-12\mu_5\mu_7-\mu_6^2+36\mu_2\mu_5^2}{N},$$

$$\mu_{11}\{\hat{\mu}_2,\hat{\mu}_4\}=\frac{\mu_6-4\mu_3^2-\mu_2\mu_4}{N},$$

$$\mu_{11}\{\hat{\mu}_2,\hat{\mu}_6\}=\frac{\mu_8-6\mu_3\mu_5-\mu_2\mu_6}{N}, \tag{25}$$

$$\mu_{11}\{\hat{\mu}_3,\hat{\mu}_4\}=\frac{\mu_7-5\mu_3\mu_4-3\mu_2\mu_5+12\mu_2^2\mu_3}{N},$$

$$\mu_{11}\{\hat{\mu}_3,\hat{\mu}_6\}=\frac{\mu_9-6\mu_4\mu_5-3\mu_2\mu_7-\mu_3\mu_6+18\mu_2^2\mu_5}{N},$$

$$\mu_{11}\{\hat{\mu}_4, \hat{\mu}_6\} = \frac{\mu_{10} - 6\mu_5^2 - 4\mu_3\mu_7 - \mu_4\mu_6 + 24\mu_2\mu_3\mu_5}{N}.$$

Подставляя полученные выражения (12), (13), (15), (24), (25) в формулу (23) с учетом (18), после упрощения получаем

$$\mathbf{D}\{\hat{\gamma}_6\} = c_6 / N, \tag{26}$$

где

$$c_6 = 900M_3^4M_4 + 7100M_3^4 - 180M_3^2M_4M_6 - 1180M_3^2M_6 + 9M_4M_6^2 + 86M_6^2 + \\ +1800M_3^2M_4^2 + 2100M_3^2M_4 - 180M_4^2M_6 - 225M_4^2 - 840M_4M_6 + 900M_4^3 + \\ +14400M_3^2 + 225M_8 - 5040M_3M_5 + 60M_3^2M_8 - 12M_5M_7 + 216M_5^2 - \\ -1560M_3^3M_5 - 1320M_3M_4M_5 + 156M_3M_5M_6 + M_{12} + 60M_4M_8 - \\ -6M_6M_8 + 840M_3M_7 - 40M_3M_9 - 30M_{10}. \tag{27}$$

Из выражения (26) следует, что оценка (2) является состоятельной.

Для симметричных распределений выражения (22) и (27) с учетом (18) принимают вид

$$\mathbf{M}\{\hat{\gamma}_6\} = M_6 - 15M_4 + 30,$$

$$c_6 = 9M_4M_6^2 + 86M_6^2 - 180M_4^2M_6 - 225M_4^2 - 840M_4M_6 + \\ +900M_4^3 + 225M_8 + M_{12} + 60M_4M_8 - 6M_6M_8 - 30M_{10}. \tag{28}$$

Ошибки оценок кумулянтных коэффициентов типовых распределений. Применим полученные результаты для анализа ошибок оценок коэффициентов γ_s, $s = \overline{3,6}$, нескольких типовых симметричных распределений [17] (табл. 1). Для анализа ошибок оценок коэффициентов γ_3 и γ_4 воспользуемся результатами работы [5]. Оценки коэффициентов γ_3 и γ_4 находим по формулам

$$\hat{\gamma}_3 = \frac{\hat{\mu}_3}{(\hat{\mu}_2)^{3/2}}, \quad \hat{\gamma}_4 = \frac{\hat{\mu}_4}{(\hat{\mu}_2)^2} - 3.$$

Их математические ожидания и дисперсии записываем в виде [5] $\mathbf{M}\{\hat{\gamma}_3\} = \gamma_3$, $\mathbf{M}\{\hat{\gamma}_4\} = \gamma_4$, $\mathbf{D}\{\hat{\gamma}_3\} = c_3 / N$, $\mathbf{D}\{\hat{\gamma}_4\} = c_4 / N$, где c_3, c_4 — коэффициенты, вычисляемые по формулам

$$c_3 = M_6 - 3M_3M_5 - 6M_4 + \frac{9}{4}M_3^2M_4 + \frac{35}{4}M_3^2 + 9, \tag{29}$$

$$c_4 = M_8 - 4M_4M_6 - 8M_3M_5 + 4M_4^3 - M_4^2 + 16M_3^2M_4 + 16M_3^2, \tag{30}$$

а моменты M_s — по формуле (18). Оценки коэффициентов γ_s, $s=\overline{3,6}$, являются несмещенными, поэтому их среднеквадратические ошибки имеют вид $\varepsilon_s=\sqrt{\mathbf{D}\{\hat{\gamma}_s\}}=\sqrt{c_s/N}$ и при фиксированном объеме выборки зависят только от коэффициентов c_s.

Значения коэффициентов c_s оценок $\hat{\gamma}_s$, $s=\overline{3,6}$, рассчитанные по формулам (21), (28)—(30), приведены в табл. 2, из которой видно, что при фиксированном объеме N выборки среднеквадратические ошибки оценок возрастают с возрастанием порядка оцениваемого коэффициента и существенно зависят от типа распределения.

Таблица 1

Распределение	Плотность вероятности	Центральный момент
Гауссовское	$p(x)=\dfrac{1}{\sigma\sqrt{2\pi}}\exp\left\{-\dfrac{(x-m)^2}{2\sigma^2}\right\}$	$\mu_s=\begin{cases}0, s-\text{нечетное;}\\ \sigma^s(s-1)(s-3)\ldots 3\cdot 1, s-\text{четное.}\end{cases}$
Лапласа	$p(x)=\dfrac{1}{2}\lambda\exp(-\lambda\|x-m\|)$	$\mu_s=\begin{cases}0, s-\text{нечетное;}\\ \dfrac{s!}{\lambda^s}, s-\text{четное.}\end{cases}$
Стьюдента	$p(x)=\dfrac{\Gamma\left(\dfrac{\nu+1}{2}\right)}{\sqrt{\pi\nu}\Gamma\left(\dfrac{\nu}{2}\right)}\left(1+\dfrac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}$, где $\Gamma(x)$ — гамма-функция	$\mu_s=\begin{cases}0, s-\text{нечетное;}\\ \dfrac{1\cdot 3\cdot 5\cdot\ldots\cdot(s-1)\nu^{s/2}}{(\nu-2)(\nu-4)\ldots(\nu-s)}, s-\text{четное,}\end{cases}$ $\nu>s$
Логистическое	$p(x)=\dfrac{\exp\left(\dfrac{x-m}{\lambda}\right)}{\lambda\left[1+\exp\left(\dfrac{x-m}{\lambda}\right)\right]^2}$	$\mu_s=\begin{cases}0, s-\text{нечетное;}\\ (\lambda\pi)^s(2^s-2)\|B_s\|, s-\text{четное,}\end{cases}$ где B_s – числа Бернулли, $B_2=1/6$, $B_4=-1/30$, $B_6=1/42$, $B_8=-1/30$, $B_{10}=5/66$, $B_{12}=-691/2730$

Таблица 2

Распределение	Значение коэффициента			
	c_3	c_4	c_5	c_6
Гауссовское	6	24	120	720
Лапласа	63	$1{,}188\cdot 10^3$	$7{,}2\cdot 10^4$	$4{,}479\cdot 10^6$
Стьюдента, $\nu=13$	15,810	204,868	$8{,}601\cdot 10^3$	$1{,}435\cdot 10^6$
Стьюдента, $\nu=200$	6,282	26,317	140,555	918,077
Логистическое	23,657	294,912	$9{,}126\cdot 10^3$	$3{,}463\cdot 10^5$

Для сравнения в табл. 3 приведены рассчитанные среднеквадратические ошибки оценок кумулянтных коэффициентов γ_s для выборок объема $N = 10^6$. Из табл. 3 видно, что наибольшие среднеквадратические ошибки имеют оценки кумулянтных коэффициентов распределения Лапласа, наименьшие — гауссовского распределения.

Полученные результаты позволяют определить диапазон возможных значений оценок кумулянтных коэффициентов. Если использовать результаты работы [5] и предположить, что оценки имеют асимптотически гауссовское распределение, то значения оценок лежат в диапазоне $\mathbf{M}\{\hat{\gamma}_s\} \pm 3\varepsilon_s$ с вероятностью 0,9973.

Среднеквадратические ошибки характеризуют величину абсолютного отклонения результатов оценивания кумулянтных коэффициентов γ_s от их истинных значений. Поэтому целесообразно сравнить при фиксированном объеме N выборки относительные ошибки оценок кумулянтных коэффициентов $\gamma_s \neq 0$, которые в данном случае имеют вид

$$\delta_s = \frac{\varepsilon_s}{\gamma_s} = \frac{\sqrt{\mathbf{D}\{\hat{\gamma}_s\}}}{\gamma_s} = \frac{1}{\gamma_s}\sqrt{\frac{c_s}{N}}. \tag{31}$$

В табл. 4 приведены теоретические значения кумулянтных коэффициентов γ_s и рассчитанные относительные ошибки их оценок для выборок объема $N = 10^6$. Из табл. 4 видно, что для оценки коэффициента γ_4 всех

Таблица 3

Распределение	Значение среднеквадратической ошибки			
	ε_3	ε_4	ε_5	ε_6
Гауссовское	0,002	0,005	0,011	0,027
Лапласа	0,008	0,034	0,268	2,117
Стьюдента, $\nu = 13$	0,004	0,014	0,093	1,198
Логистическое	0,005	0,017	0,096	0,588

Таблица 4

Распределение	γ_s	δ_s, %
Лапласа	$\gamma_4 = 3$	1,133
	$\gamma_6 = 30$	7,057
Стьюдента, $\nu = 13$	$\gamma_4 = 0{,}6667$	2,100
	$\gamma_6 = 3{,}8095$	31,448
Логистическое	$\gamma_4 = 1{,}2$	1,417
	$\gamma_6 = 6{,}8571$	8,575

Таблица 5

Распределение	γ_s	$N_{\min}$ при δ_s	
		0,01	0,05
Лапласа	$\gamma_4 = 3$	$1{,}32 \cdot 10^6$	$0{,}053 \cdot 10^6$
	$\gamma_6 = 30$	$49{,}77 \cdot 10^6$	$1{,}991 \cdot 10^6$
Стьюдента, $\nu = 13$	$\gamma_4 = 0{,}6667$	$4{,}610 \cdot 10^6$	$0{,}184 \cdot 10^6$
	$\gamma_6 = 3{,}8095$	$988{,}618 \cdot 10^6$	$39{,}545 \cdot 10^6$
Стьюдента, $\nu = 200$	$\gamma_4 = 0{,}0306$	$2{,}808 \cdot 10^8$	$0{,}112 \cdot 10^8$
	$\gamma_6 = 0{,}0063$	$2304{,}477 \cdot 10^8$	$92{,}179 \cdot 10^8$
Логистическое	$\gamma_4 = 1{,}2$	$2{,}048 \cdot 10^6$	$0{,}082 \cdot 10^6$
	$\gamma_6 = 6{,}8571$	$73{,}655 \cdot 10^6$	$2{,}946 \cdot 10^6$

рассмотренных распределений можно ограничиться объемом выборки $N = 10^6$, для оценки коэффициента γ_6 его недостаточно.

Используя формулу (31), можно определить минимальный объем выборки $N_{\min}$, при котором обеспечивается заданная относительная ошибка δ_s: $N_{\min} = c_s / (\delta_s \gamma_s)^2$. В табл. 5 приведены значения объема выборки $N_{\min}$, обеспечивающие относительную ошибку оценок $\delta_s = 0{,}01$ (1%) и $\delta_s = 0{,}05$ (5%) кумулянтных коэффициентов γ_4 и γ_6.

Результаты, приведенные в табл. 5, свидетельствуют о том, что для распределений, кумулянтные коэффициенты которых существенно отличаются от нуля (Лапласа, Стьюдента ($\nu = 13$), логистического), порядок минимальных объемов выборки $N_{\min}$ для обеспечения относительных ошибок 1% и 5% для γ_4 составляет соответственно 10^6 и 10^5, а для γ_6 — 10^7 и 10^6. При стремлении значений кумулянтных коэффициентов к нулю (распределение Стьюдента, $\nu = 200$) необходимо существенное увеличение минимального объема выборки: для γ_4 — на два порядка, для γ_6 — на четыре порядка.

Выводы

Полученные формулы позволяют вычислять среднеквадратическую ошибку оценок кумулянтных коэффициентов пятого и шестого порядков по известным центральным моментам.

На основе сравнительного анализа среднеквадратических ошибок оценок кумулянтных коэффициентов типовых симметричных распределений установлено, что для оценок кумулянтных коэффициентов γ_4 и γ_6, значения которых существенно отличаются от нуля, относительная ошибка 5 % обеспечивается соответственно при объеме выборки не меньше 10^5 и 10^6. При стремлении значений кумулянтных коэффициентов к нулю минимальный объем выборки возрастает на несколько порядков.

General expressions have been obtained for determining mathematical expectation and dispersion of estimates for cumulant coefficients of the fifth and sixth orders which permit calculating mean-root-square errors of estimates of these coefficients. Estimation errors of cumulant coefficients γ_s, $s=\overline{3,6}$, some type symmetric distributions at the fixed volume of a sample have been analyzed. Minimum volume of the sample is determined, when a preset relative error of estimates of cumulant coefficients of these distributions is provided.

СПИСОК ЛИТЕРАТУРЫ

1. *Левин Б.Р., Шварц В.* Вероятностные модели и методы в системах связи и управления. — М. : Радио и связь, 1985. — 312 с.
2. *Новицкий П.В., Зограф И.А.* Оценка погрешностей результатов измерений. — Л. : Энергоатомиздат, 1991. — 304 с.
3. *Шелухин О.И., Беляков И.В.* Негауссовские процессы. — СПб. : Политехника, 1992. — 312 с.
4. *Кендалл М., Стьюарт А.* Теория распределений / Пер. с англ. В.В. Сазонова, А.Н. Ширяева; под ред. А.Н. Колмогорова. — М. : Наука, 1966. — 588 с.
5. *Крамер Г.* Математические методы статистики / Пер. с англ. А.С. Монина и А.А. Петрова; под ред. А.Н. Колмогорова. — М. : Мир, 1975. — 648 с.
6. *Митропольский А.К.* Техника статистических вычислений. — М. : Наука, 1971. — 576 с.
7. *Хан Г., Шапиро С.* Статистические модели в инженерных задачах / Пер. с англ. Е.Г. Коваленко; под ред. В.В. Налимова. — М. : Мир, 1969. — 396 с.
8. *Марченко Б.Г., Мыслович М.В.* Вибродиагностика подшипниковых узлов электрических машин. — К. : Наук. думка, 1992. — 196 с.
9. *Берегун В.С., Красильніков О.І.* Дослідження областей невід'ємності ортогональних подань щільності імовірностей // Электроника и связь. — 2010. — № 3 (56). — С. 73—78.
10. *Красильников А.И., Пилипенко К.П.* Применение двухкомпонентной гауссовской смеси для идентификации одновершинных симметричных плотностей вероятностей // Там же. — 2008. — № 5 (46). — С. 20—29.
11. *Гармаш О.В., Красильников А.И.* Применение функций Пирсона для аппроксимации пуассоновской спектральной плотности Колмогорова линейных случайных процессов // Електроніка та системи управління. — 2011. — № 3 (29). — С. 50—59.
12. *Берегун В.С., Горовецька Т.А., Красильніков О.І.* Статистичний аналіз шумів колінних суглобів // Акустичний вісник. — 2011. — **14**, № 2. — С. 3—15.
13. *Айвазян С.А., Енюков И.С., Мешалкин Л.Д.* Прикладная статистика: Основы моделирования и первичная обработка данных.— М. : Финансы и статистика, 1983. — 471 с.
14. *Кунченко Ю.П.* Полиномиальные оценки параметров близких к гауссовским случайных величин. Ч. I. Стохастические полиномы, их свойства и применение для нахождения оценок параметров. — Черкассы: ЧИТИ, 2001. — 133 с.
15. *Бендат Дж., Пирсол А.* Прикладной анализ случайных данных / Пер с англ. В.Е. Привальского и А.И. Кочубинского; под ред. И.Н. Коваленко. — М. : Мир, 1989. — 540 с.
16. *Тихонов В.И.* Статистическая радиотехника. — М. : Радио и связь, 1982. — 624 с.
17. *Вадзинский Р.Н.* Справочник по вероятностным распределениям. — СПб. : Наука, 2001. — 295 с.

Поступила 22.10.13;
после доработки 04.12.13

БЕРЕГУН Виктор Сергеевич, канд. техн. наук, ассистент кафедры акустики и акустоэлектроники факультета электроники Национального технического университета Украины «Киевский политехнический ин-т», который окончил в 2004 г. Область научных исследований — вероятностные характеристики и методы обработки негауссовских сигналов и их функциональных преобразований в акустических информационных системах.

ГАРМАШ Оксана Викторовна, ассистент кафедры акустики и акустоэлектроники факультета электроники Национального технического университета Украины «Киевский политехнический ин-т», который окончила в 2005 г. Область научных исследований — теоретические и экспериментальные исследования вероятностных характеристик флуктуационных процессов в акустических информационных системах.

КРАСИЛЬНИКОВ Александр Иванович, канд. физ.-мат. наук, доцент, ст. науч. сотр. Ин-та технической теплофизики НАН Украины. В 1973 г. окончил Киевский политехнический ин-т. Область научных исследований — математические модели, вероятностные характеристики и методы статистической обработки флуктуационных сигналов в системах шумовой диагностики.