

© Сухорученков Б. И.
Sukhoruchenkov B.

СПОСОБ ПОВЫШЕНИЯ ТОЧНОСТИ ОЦЕНИВАНИЯ НОРМАЛЬНОГО РАСПРЕДЕЛЕНИЯ ХАРАКТЕРИСТИК ТЕХНИЧЕСКИХ СИСТЕМ ПО ОГРАНИЧЕННОЙ ВЫБОРКЕ

METHOD OF INCREASE OF EXACTNESS OF EVALUATION OF NORMAL DISTRIBUTION OF DESCRIPTIONS OF TECHNICAL SYSTEMS ON A LIMIT SELECTION

Аннотация. Рассмотрены методы статистического оценивания случайных характеристик технических систем (ТС) с нормальным распределением, плотность вероятности которого может быть представлена в виде зависимости от среднеквадратического отклонения (СКО) или от показателя компактности распределения, обратного СКО. Показано, что при втором варианте достигается более высокая точность оценивания характеристик ТС на основе ограниченной выборки по методу несмещенных оценок.

Annotation. The methods of statistical evaluation of casual descriptions of the technical systems (TC) are considered with normal distribution, the closeness of probability of that can be presented as dependence on a root-mean-square rejection (RMSR) or from the index of compactness of distribution, reverse RMSR. It is shown that at the second variant higher exactness of evaluation of descriptions TC is arrived at on the basis of a limit selection on the method of the undisplaced estimations.

Ключевые слова. Случайная характеристика, нормальное распределение, параметры распределения, выборка, статистический метод, точечная оценка, среднеквадратическое отклонение.

Key words. Casual description, normal distribution, parameters of distribution, selection, statistical method, point estimation, root-mean-square rejection.

1. Модели представления нормального распределения характеристик технических систем

Характеристики X технических систем (ТС) и их подсистем (кинематические, энергетические, прочностные, электрические и др.) обычно являются случайными и имеют нормальное распределение (НР). При этом плотность вероятности (ПВ) характеристик представляется в виде зависимости от математического ожидания (МО) M и среднеквадратического отклонения (СКО) σ или дисперсии $D=\sigma^2$, которые являются мерой рассеивания $f_1(x, M, \sigma) = (2\pi)^{-0.5} \sigma^{-1} \exp[-0,5 \sigma^{-2}(x - M)^2]$, $\sigma > 0$. (1)

Модель ПВ характеристик можно также представить в функции от меры точности или сосредоточенности распределения, обратного СКО [1, 3]. Введем в рассмотрение показатель компактности (кучности) распределения (ПК) $G=\sigma^{-1}$. При этом модель ПВ нормального распределения характеристик ТС записывается в виде

$$f_2(x, M, G) = (2\pi)^{-0.5} G \exp[-0,5 G^2(x - M)^2], G > 0. \quad (2)$$

Модели (1) и (2) представления ПВ случайных характеристик ТС (далее просто модели) при $G=\sigma^{-1}$ совпадают. Однако при неизвестных параметрах распределения приводят к разным результатам оценивания характеристик ТС по экспериментальным данным, что показано ниже.

2. Классические методы оценивания параметров распределения характеристик ТС по экспериментальным данным

Если параметры распределения характеристик ТС не известны, они оцениваются на основе модели (1) по данным независимых наблюдений реализаций характеристик x_i , $i = 1, \dots, n$. Для этого используются классические методы наименьших квадратов, максимального правдоподобия и др. [1–5]. В соответствии с этими методами сначала определяются точечные оценки МО и дисперсии

Сухорученков Борис Иванович – доктор технических наук, профессор, член-корреспондент Российской инженерной академии, профессор кафедры баллистических ракет, Военная академия РВСН имени Петра Великого, тел. (495) 696-06-48.

Sukhoruchenkov Boris – doctor of technical sciences, professor, corresponding member of the Russian engineering Academy, professor of the department of ballistic missiles, the Military Academy of strategic missile forces named after Peter the Great, tel. (495) 696-06-48.

$D=\sigma^2$ и их дисперсии, а оценки СКО вычисляются приближенно на основе метода линеаризации. Для отличия этих оценок от оценок по другим методам обозначим их индексом «1»:

$$\bar{M}_1 = \frac{1}{n} \sum_{i=1}^n x_i; \quad \sigma_{\bar{M}_1}^2 = \frac{1}{n} \bar{D}_1; \quad (3)$$

$$\bar{D}_1 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{M})^2; \quad \sigma_{\bar{D}_1}^2 = \frac{2}{n} \bar{D}_1^2; \quad (4)$$

$$\bar{\sigma}_1 = \sqrt{\bar{D}}; \quad \sigma_{\bar{\sigma}_1}^2 = \frac{2}{n} \bar{D}_1. \quad (5)$$

Интервальные оценки для МО определяются на основе распределения Стьюдента, а для дисперсии – на основе χ^2 – распределения [1–4]. Оценки (3)–(5) являются асимптотически несмещенными и асимптотически эффективными (при $n \rightarrow \infty$). Однако при ограниченной выборке оценки могут иметь смещение и терять эффективность.

Пример 1. Получены результаты наблюдений значений характеристики ТС $\{x_i\} = \{10,7; 14,1; 7,4; 8,2; 11,5; 5,7; 10,3\}$. Предполагается, что характеристика X имеет нормальное распределение. Необходимо оценить параметры распределения классическими методами.

На основе полученной выборки по зависимостям (3)–(5) получаются реализации оценок МО, СКО и СКО оценок параметров модели (1), приведенные в табл. 1.

Реализации оценок МО, СКО и СКО – оценок параметров модели (1)

Метод оценивания	Модель ПВ	Зависимости	Реализации оценок параметров и их СКО					
			\hat{M}	$\sigma_{\hat{M}}$	$\hat{\sigma}$	$\sigma_{\hat{\sigma}}$	\hat{G}	$\sigma_{\hat{G}}$
Классические методы	(1)	(3)–(5) (12), (13)	9,70	1,06	2,81	0,75	-	-
				1,47		1,72		
МНО	(1)	(8)–(11)	9,70	1,47	3,67	1,50	-	-
	(2)	(8)–(11) (14), (16)	9,70	1,16	-	-	0,370	0,101

3. Оценивание параметров распределения характеристик ТС по методу несмещенных оценок

3.1. Построение плотности вероятности оценок параметров распределения характеристик ТС

В работе [6] обоснован метод несмещенных оценок (МНО), обеспечивающий получение несмещенных и эффективных оценок параметров распределения характеристик ТС даже по ограниченной выборке. В соответствии с МНО сначала строится ПВ возможных оценок параметров распределения. Для сокращения записей обозначим параметры σ и G в общем виде R . При этом для разных моделей ПВ распределения характеристик ТС введем обозначение $f(x, M, R)$. Возможные оценки параметров M, σ, G и R обозначим в виде m, s, g и r соответственно. При принятых обозначениях ПВ оценок пара-

метров M и R согласно МНО строится последовательно по зависимостям

$$\left. \begin{aligned} q(m, r) &= \prod_{i=1}^n f(x_i, m, r); \\ k &= \int_0^\infty \int_{-\infty}^\infty q(m, r) dm dr; \\ f(m, r) &= k^{-1} q(m, r), \end{aligned} \right\} \quad (6)$$

где $f(x_i, m, r)$ – ПВ (1) или (2) при значениях $\{x_i\}$ и замене МО M и соответствующего параметра R на их возможные оценки m и r .

На основе ПВ $f(m, r)$ можно построить автономные ПВ оценок параметров НР

$$f(m) = \int_0^\infty f(m, r) dr; \quad f(r) = \int_{-\infty}^\infty f(m, r) dm. \quad (7)$$

В зависимости от модели ПВ характеристик ТС по формулам (6) и (7) получаются ПВ $f(m, s)$ оценок параметров МО и СКО или ПВ $f(m, g)$ оценки МО и ПК и автономные ПВ $f(m), f(s)$ или $f(g)$. ПВ (6) и (7) являются исчерпывающей характеристикой оценок параметров НР характеристик ТС как случайных векторов и случайных величин.

3.2. Оценивание параметров распределения характеристик ТС

На основе построенных ПВ оценок параметров НР можно осуществить как точечное, так и интервальное

оценивание неизвестных параметров. Точечные оценки параметров НР и их дисперсии с учетом принятых обозначений определяются на основе автономных ПВ по зависимостям

$$\bar{M} = \int_{-\infty}^\infty m f(m) dm; \quad (8)$$

$$\sigma_{\bar{M}}^2 = \int_{-\infty}^\infty (m - \bar{M})^2 f(m) dm; \quad (9)$$

$$\bar{R} = \int_0^\infty r f(r) dr; \quad (10)$$

$$\sigma_{\bar{R}}^2 = \int_0^\infty (r - \bar{R})^2 f(r) dr. \quad (11)$$

В зависимости от модели ПВ по формулам (8)–(11) определяются точечные оценки МО M , параметров σ или G и их дисперсии. Интервальные оценки неизвест-

ных параметров распределения вычисляются численным способом на основе интегрирования автономных ПВ (7).

Получаемые оценки (8)–(11) являются несмещенными и эффективными даже при ограниченной выборке.

Заметим, что на основе построенной ПВ оценок $f(m, s)$ можно уточнить дисперсии и СКО точечных оценок параметров распределения, приведенных в п.2, по зависимостям (9) и (11) при подстановке в них оценок параметров M и σ , полученных классическими методами

$$\sigma_{\bar{M}_1}^2 = \int_{-\infty}^{\infty} (m - \bar{M}_1)^2 f(m) dm; \quad (12)$$

$$\sigma_{\bar{\sigma}_1}^2 = \int_0^{\infty} (s - \bar{\sigma}_1)^2 f(s) ds. \quad (13)$$

3.3. Связь оценок СКО по моделям (1) и (2) ПВ распределения характеристик ТС

Если получена оценка ПВ ПК $f(g)$ по модели (2), то можно оценить СКО распределения характеристик ТС по модели (1) как функции от ПК. Возможная оценка СКО s связана с оценкой g : $s = 1/g$. Поэтому точечная оценка СКО и ее дисперсия определяются как функция от случайного аргумента по зависимостям

$$\bar{\sigma} = \int_0^{\infty} (1/g) \cdot f(g) dg; \quad (14)$$

$$\sigma_{\bar{\sigma}}^2 = \int_0^{\infty} (1/g - \bar{\sigma})^2 \cdot f(g) dg.$$

Кроме того, для оценивания СКО можно построить автономную ПВ $f(s)$ оценок СКО s на основе ПВ оценок g , которую для этого обозначим в виде $f_2(g)$. По методике построения ПВ функции от случайного аргумента [1, 3] с учетом связи оценок $s = 1/g$ после преобразований получим

$$f(s) = s^{-2} f_2(s^{-1}). \quad (15)$$

На основе ПВ (15) определяется точечная оценка СКО характеристик ТС и ее дисперсия по зависимостям

$$\bar{\sigma} = \int_0^{\infty} s f(s) ds; \quad \sigma_{\bar{\sigma}}^2 = \int_0^{\infty} (s - \bar{\sigma})^2 f(s) ds. \quad (16)$$

Оценки (14) и (16) совпадают, при этом на основе ПВ (15) можно осуществить интервальное оценивание СКО характеристик ТС при использовании модели (2).

Пример 2. В условиях примера 1 оценить параметры распределения характеристики ТС по МНО на основе модели (1) и сравнить их с оценками параметров классическими методами.

Вычисления и построение графиков здесь и далее проводились в системе MathCAD. Незвестными параметрами модели (1) ПВ распределения являются МО и СКО. ПВ оценок этих параметров, построенные по зависимостям (6) и (7), показаны на рис. 1 и 2. Из рис. 2 видно, что

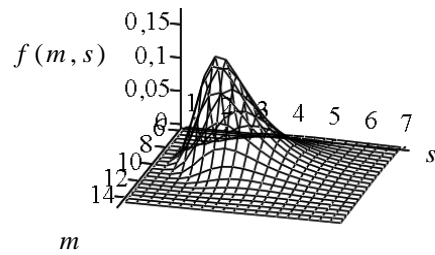


Рис. 1. Общий вид ПВ оценок параметров $f(m, s)$

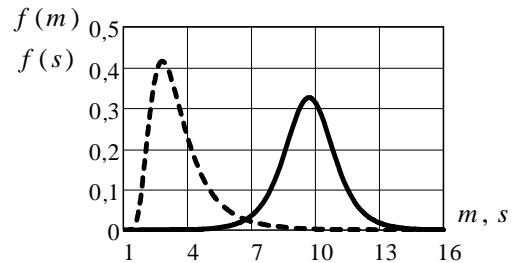


Рис. 2. Автономные ПВ оценок параметров $f(m)$ (сплошная) и $f(s)$ (пунктирная)

ПВ оценок СКО имеет удлинённый «хвост» при повышенных значениях оценок.

Реализации точечных оценок параметров распределения характеристики ТС и их СКО, вычисленные по зависимостям (8)–(11) при ПВ оценок, показанных на рис. 2, представлены в таблице. Здесь же приведены СКО оценок параметров, полученных классическими методами, по зависимостям (12) и (13). Видно, что оценка МО по МНО совпадает с оценкой классическими методами, однако СКО оценки МО, получаемая классическими методами, занижена. Оценка СКО по классическим методам также занижена и имеет более высокую погрешность.

Пример 3. В условиях примера 1 оценить параметры распределения характеристик ТС по МНО на основе модели (2) и сравнить их с оценками параметров модели (1).

При модели (2) неизвестными параметрами распределения являются МО и ПК G . ПВ оценок этих параметров, построенные по зависимостям (6) и (7), показаны на рис. 3–5. На рис. 4 для сравнения приведена также ПВ оценок МО по модели (1). На основе ПВ оценок ПК $f(g)$ строится ПВ оценок СКО по зависимости (15), которая показана на рис. 6. Здесь же показана ПВ оценок СКО при модели (1). Из рис. 4 и 6 видно, что ПВ оценок МО и СКО по модели (2) имеют более компактный вид.

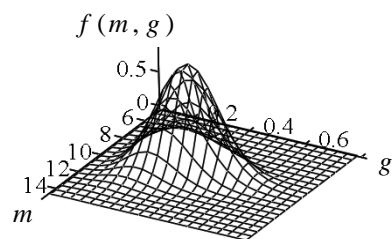


Рис. 3. Общий вид ПВ оценок параметров $f(mg)$

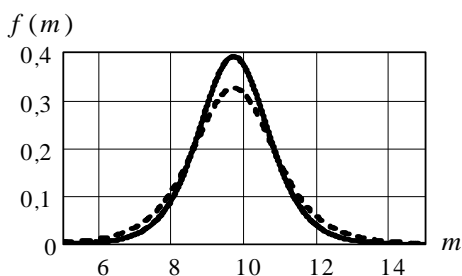


Рис. 4. ПВ оценок МО $f(m)$ по модели (2) (сплошная) и по модели (1) характеристики ТС (пунктирная)

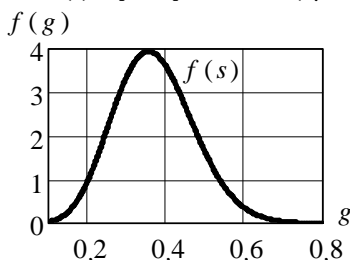


Рис. 5. ПВ оценок параметра компактности $f(g)$

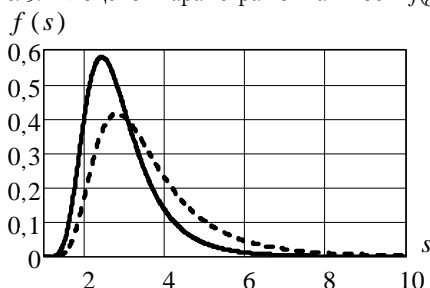


Рис. 6. ПВ оценок СКО по зависимости (15) (сплошная) и по модели (1) характеристики ТС (пунктирная)

Реализации оценок параметров модели (2) и их СКО, вычисленные по зависимостям (8)–(11), приведены в таблице. Характерно, что оценки МО характеристики ТС по модели (1) и (2) совпадают, однако по модели (2) получается более высокая точность оценок МО. Оценка СКО, полученная по зависимости (14) или (16), меньше оценки СКО по модели (1) и имеет более высокую точность.

Заметим, что несмещенные точечные оценки параметров σ и G , реализации которых приведены в таблице в последней строке, не соответствуют и не должны соответствовать равенству $G = \sigma^{-1}$, которое справедливо только для известных параметров распределения.

На основе данных таблицы следует, что при использовании модели (2) достигается более высокая точность оценивания параметров распределения характеристик ТС.

4. Оценивание плотности вероятности распределения характеристик ТС

Для анализа возможных вариаций характеристик ТС необходимо оценить ПВ распределения характеристик $f(x)$. Точечная оценка ПВ характеристик ТС оперативно определяется по методу линеаризации на основе

зависимости (1) или (2) при полученных точечных оценках параметров распределения. Кроме того, на основе МНО с учетом ПВ оценок параметров $f(m, r)$ также можно построить несмещенную точечную оценку ПВ распределения характеристик ТС и ее дисперсию по зависимостям

$$\bar{f}(x) = \int_0^{\infty} \int_{-\infty}^{\infty} f(x, m, r) f(m, r) dm dr; \quad (17)$$

$$\sigma_{\bar{f}(x)}^2 = \int_0^{\infty} \int_{-\infty}^{\infty} [f(x, m, r) - \bar{f}(x)]^2 f(m, r) dm dr, \quad (18)$$

где $f_2(x, m, r)$ – ПВ (1) или (2), в которой неизвестные параметры M и R заменяются на их возможные оценки m и r .

Точечная оценка ПВ (17) практически совпадает с ПВ (1) или (2), в которой вместо неизвестных параметров распределения используются их точечные оценки. Однако для построения ПВ по зависимости (17) требуется большой объем вычислений.

Пример 4. В условиях примеров 1 и 2 построить оценку ПВ характеристики ТС по модели (1) по МНО.

Реализация оценки ПВ характеристики ТС, построенная по зависимости (17) при ПВ $f(m, s)$, показанной на рис. 1, а также оценка, полученная оперативно по зависимости (1) при оценках параметров $\hat{M}=9,70$ и $\hat{\sigma}=3,67$ см. таблицу, показана на рис. 7. Из рис. 7 следует, что оценки ПВ, построенные оперативно и по зависимости (17), практически совпадают. При этом более точная оценка ПВ по зависимости (17) имеет удлиненные «хвосты».

Пример 5. В условиях примеров 1, 2 и 3 построить оценки ПВ характеристики ТС различными методами и сравнить их.

Вычисления проводились оперативным методом при оценках параметров распределения характеристики ТС, полученных разными методами по моделям (1) и

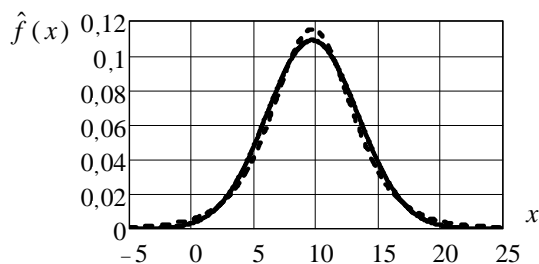


Рис. 7. Реализации оценок ПВ характеристики ТС, полученные оперативно (сплошная) и по зависимости (17) (пунктирная)

(2). Реализации оценок ПВ характеристики ТС показаны на рис. 8.

Из рис. 8 видно, что оценки ПВ, полученные разными методами, близки, однако ПВ по модели (2) имеет наибольшую концентрацию.

Пример 6. В условиях примеров 1, 2, 3 и 5 определить СКО оценок ПВ характеристики ТС по моделям (1) и (2) разными методами и сравнить точность оценок ПВ.

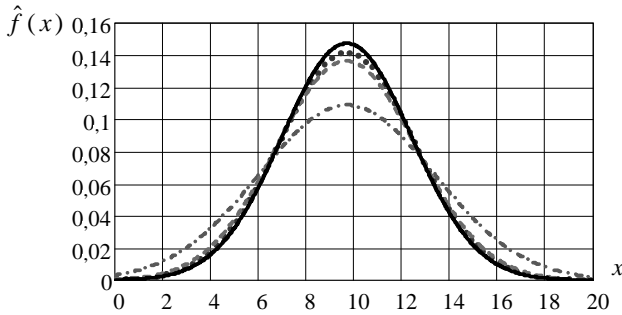


Рис. 8. Реализации оценок ПВ характеристики ТС:
 – по модели (1) при оценках параметров классическими методами; - . - . - по модели (1) при оценках параметров по МНО; - - - - по модели (1) при оценке по зависимости (14) или (16); — — — — по модели (2) при оценках параметров по МНО

Вычисления СКО проводились по зависимостям (18) на основе ПВ оценок параметров распределения, построенных по МНО, при оценках ПВ характеристики ТС, показанных на рис. 8. СКО оценок ПВ характеристики ТС показаны на рис. 9.

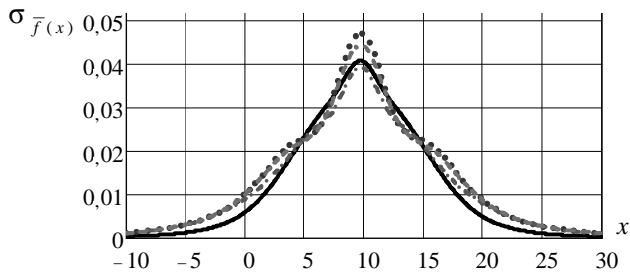


Рис. 9. СКО оценок ПВ характеристики ТС:
 – по модели (1) при оценках параметров классическими методами; - . - . - по модели (1) при оценках параметров по МНО; - - - - по модели (1) при оценке СКО по зависимости (14) или (16); — — — — по модели (2) при оценках параметров по МНО

Из рис. 9 следует, что максимальные погрешности оценок ПВ наблюдаются при характеристиках ТС в области оценки МО. Однако при анализе возможных вариаций характеристик ТС наибольшее влияние оказывают «хвосты» распределения. При этом наименьшие погрешности оценок «хвостов» распределения характеристики ТС достигаются при использовании модели (2).

Таким образом, результаты проведенных исследований показывают, что для повышения точности оценивания параметров и плотности вероятности нормального распределения характеристик ТС целесообразно использовать модель ПВ характеристик (2) и метод несмещенных оценок.

5. Зависимость эффективности модели (2) ПВ распределения характеристик ТС от объема выборки

Точность оценивания характеристик ТС зависит от модели ПВ распределения характеристик, от объема выборки и от статистического метода. Зависимости СКО оценок МО и СКО характеристик ТС по МНО от объема выборки при показателях, соответствующих примеру 1, отнесенные к оценке СКО характеристики ТС, полученной в примере 2, показаны на рис. 10.

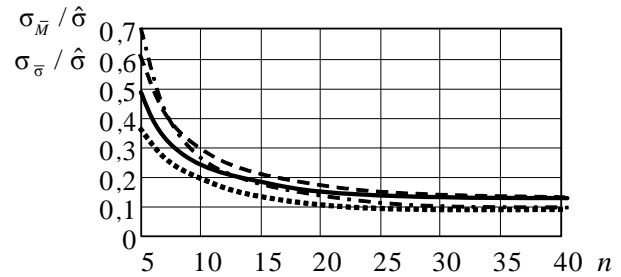


Рис. 10. Зависимости СКО оценок МО и СКО характеристики ТС, отнесенные к оценке СКО $\hat{\sigma}$, от объема выборки n :
 - - - - – СКО оценки МО по модели (1);
 — — — — – СКО оценки МО по модели (2);
 - . - . - – СКО оценки СКО по модели (1);
 – СКО оценки СКО по модели (2)

Из рис. 10 видно, что СКО оценок МО и СКО характеристик ТС интенсивно убывают при возрастании объема выборки до $n \approx 30-40$. При этом СКО оценок параметров распределения характеристик ТС по модели (2) заметно ниже по сравнению с СКО оценок по модели (1).

Для анализа эффективности использования модели (2) ПВ характеристик ТС рассмотрим СКО оценок МО $\sigma_{\bar{M}_2}$ и $\sigma_{\bar{M}_1}$ и СКО оценок СКО характеристик ТС $\sigma_{\bar{\sigma}_2}$ и $\sigma_{\bar{\sigma}_1}$, получаемые по МНО на основе моделей (1) и (2). Зависимости отношений СКО оценок параметров модели (2) к СКО оценок параметров модели (1) от объема выборки приведены на рис. 11.

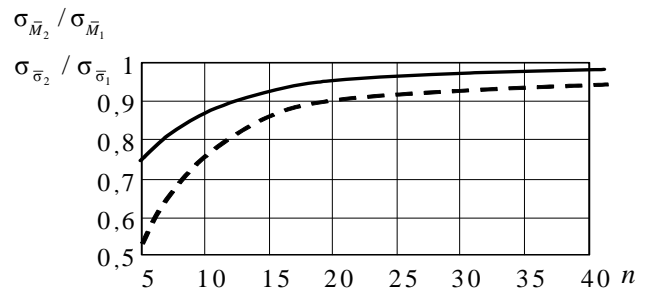


Рис. 11. Зависимости отношений СКО оценок МО и СКО распределения характеристики ТС от объема выборки n :
 — — — — – $\sigma_{\bar{M}_2} / \sigma_{\bar{M}_1}$; - - - - – $\sigma_{\bar{\sigma}_2} / \sigma_{\bar{\sigma}_1}$

Из рис. 11 следует, что точность оценивания характеристик ТС в виде модели (2) равномерно выше по сравнению с моделью (1). При малой выборке $n \approx 5-10$ на

основе модели (2) достигается повышение точности оценок МО и СКО характеристик ТС по МНО до 15-45%. При возрастании объема выборки эффективность модели (2) по сравнению с моделью (1) снижается и при $n > 50-100$ СКО оценок характеристик ТС уменьшается не более, чем на 2-5%.

6. Выводы

Проведены исследования методов статистического оценивания нормально распределенных характеристик ТС, плотности вероятности которых могут представляться в виде моделей (1) или (2) как функции от СКО или от показателя концентрации распределения.

На основе метода несмещенных оценок получены зависимости для построения плотности вероятности оценок неизвестных параметров распределения и зависимости для получения несмещенных точечных оценок

параметров распределения и их дисперсий по экспериментальным данным при использовании моделей (1) или (2) плотности вероятности характеристик ТС.

На примерах показано, что метод несмещенных оценок имеет более высокую точность оценивания характеристик ТС по сравнению с классическими методами. При этом по методу несмещенных оценок на основе модели (2) получаются более точные оценки неизвестных параметров распределения и плотности вероятности распределения характеристик ТС, особенно при малой выборке.

Таким образом, случайные характеристики ТС с нормальным распределением целесообразно оценивать по методу несмещенных оценок на основе модели (2) плотности вероятности нормального распределения характеристик ТС.

Литература

1. Венцель Е. С. Теория вероятностей. М.: «Академия», 2003.
2. Вероятность и математическая статистика: Энциклопедия / под ред. Ю. В. Прохорова. – М.: Большая Российская энциклопедия, 2003.
3. Пугачев В. С. Теория вероятностей и математическая статистика. М.: Физматлит, 2002.
4. Смирнов Н. В., Дунин-Барковский И. В. Курс теории вероятностей и математической статистики для технических приложений. М.: Наука, 1965.
5. Сухорученков Б. И., Меньшиков В. А. Методы анализа характеристик летательных аппаратов. М.: Машиностроение, 1995.
6. Сухорученков Б. И. Анализ малой выборки. Прикладные статистические методы. М.: Вузовская книга, 2010.

Материал поступил в редакцию 12. 12. 2012 г.