УДК 519.2: 528.1
DOI 10.33764/2618-981X-2022-3-198-204

Вероятности усеченных законов распределения

В. Е. Мизин[1], А. А. Ильин[1]*

[1] Сибирский государственный университет геосистем и технологий, г. Новосибирск, Российская Федерация
* e-mail: ssga221@mail.ru

Аннотация. Предельные ошибки измерений в геодезии устанавливают в предположении о нормальном законе распределения случайных ошибок. Так как, ошибки, превышающие заданные пределы, отбраковывают, распределение оставшихся случайных ошибок измерений будет усеченным. Усеченные законы изучены недостаточно. В статье приведены результаты вычисления вероятностей попадания случайных ошибок измерений в интервалы $\pm\sigma$, $\pm 2\sigma$, $\pm 2{,}5\sigma$, $\pm 3\sigma$ для усеченного нормального и, с целью сравнения, усеченного логистического законов. Выполнен анализ полученных результатов, сделаны выводы. Вероятности усеченного нормального и нормального законов отличаются незначительно. Максимальное расхождение составило 0,043. Заменять вероятности допусков нормального закона на вероятности усеченного закона не рекомендуется. Можно использовать вероятности усеченного закона только в том случае, если это позволит избежать повторных полевых измерений.

Ключевые слова: ошибка, вероятность, усеченный закон распределения

The probabilities of the cut distribution laws

V. E. Mizin, A. A. Ilin*

[1] Siberian State University of Geosystems and Technologies, Novosibirsk, Russian Federation
* e-mail: ssga221@mail.ru

Abstract. The limit measurement errors in the geodesy are determining for the normal distribution law, But the errors, exceeding the limits are throw aside and accidental errors are obey the cut law. Cut laws are not study enough. The article is leading the results of calculate the probabilities hit accidental measurements errors for intervals $\pm\sigma$, $\pm 2\sigma$, $\pm 2{,}5\sigma$, $\pm 3\sigma$ for cut normal and cut logical distribution laws. The analysis was accomplished.The probabilities of the cut normal and normal laws is different immaterial. 0,043 - maximum of the different. The normal law replace to the cut normal law is not recommending. Probabilities of the cut law may to use, when it is allowing to avoid the repetition of the field measurements.

Keywords: error, probability, cut distribution law

Введение

Одной из задач теории математической обработки геодезических измерений является изучение законов действия случайных ошибок измерений. Такими не достаточно изученными законами распределения являются усеченные законы.

В геодезии предельные ошибки измерений считают равными 3σ, $2{,}5\sigma$, 2σ, где σ – среднее квадратическое отклонение случайной величины (СВ X) [1]. Так как ошибки, превышающие заданные пределы, отбраковывают, распределение случайных ошибок измерений будет усеченным. Тем не менее, этим обстоятельством пренебрегают и, в частности для целей статистического анализа, теоретические значения вероятностей попадания случайных ошибок в заданные интервалы, как и предельные ошибки, устанавливают из предположения о нормальном законе распределения случайных ошибок измерений. Насколько обоснована такая практика, продемонстрируем следующими расчетами.

Пусть СВ X имеет в интервале от A до B плотность вероятности $f(x)$. В интервале $a \le X \le в$, где $a > A$ и $в < B$ плотность можно вычислить по формуле:

$$f_1(x) = cf(x) \tag{1}$$

Постоянная величина с определяется из уравнения:

$$\int_a^в f_1(x)\,dx = c\int_a^в f(x)\,dx = c\{F(в) - F(a)\} = 1, \tag{2}$$

$$c = \frac{1}{F(в) - F(a)}, \tag{3}$$

где $F(x)$ – функция распределения СВ X [2].

Плотность $f_1(x)$ характеризует усеченный закон распределения СВ X.

Вероятности усеченного нормального закона

Для усеченного нормального распределения

$$c = \frac{1}{F(t_в) - F(t_a)}, \tag{4}$$

где

$$F(t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{t} e^{-\frac{t^2}{2}}\,dt \tag{5}$$

- функция нормального распределения [3 – 6],

$$t = (x - M_x) / \sigma_x. \tag{6}$$

M_x - математическое ожидание СВ X.

$$\Phi(t) = \frac{1}{\sqrt{2\pi}}\int_{-t}^{t} e^{-\frac{t^2}{2}}\,dt \tag{7}$$

- интеграл вероятностей нормального закона распределения табл. 1

Таблица 1

Вероятности нормального закона

Нормальное распределение
$P\ (\lvert\Delta\rvert < \sigma) = \Phi(1) = 0{,}68269$
$P\ (\lvert\Delta\rvert < 2\sigma) = \Phi(2) = 0{,}95450$
$P\ (\lvert\Delta\rvert < 2{,}5\sigma) = \Phi(2{,}5) = 0{,}98758$
$P\ (\lvert\Delta\rvert < 3\sigma) = \Phi(3) = 0{,}99730$

Для усеченного нормального закона

$$\Phi_1(t) = c\ \Phi(\mathrm{t}). \tag{8}$$

1. При $|\Delta|_{\text{пред.}} = 3\ \sigma$.

$a = -\ 3\sigma$, $в = +\ 3\sigma$, $M_x = 0$, $F(3) = 0{,}99865$, $F(-3) = 0{,}00135$, $c = 1{,}00271$.

2. При $|\Delta|_{\text{пред.}} = 2{,}5\ \sigma$.

$a = -\ 2{,}5\ \sigma$, в $= +2{,}5\ \sigma$, $M_x = 0$.
$F(2{,}5) = 0{,}99379$, $F(-2{,}5) = 0{,}00621$, $c = 1{,}01258$.

3. При $|\Delta|_{\text{пред .}} = 2\ \sigma$

$a = -\ 2\ \sigma$, $в = +2\sigma$, $M_x = 0$. $F(2) = 0{,}97725$, $F(-2) = 0{,}02275$, $c = 1{,}04767$.

Результаты вычисления вероятностей усеченного нормального закона приведены в табл.2.

Таблица 2

Усеченное нормальное распределение

Вероятности/пределы	$\lvert\Delta\rvert_{\text{пред.}} = 3\ \sigma$.	$\lvert\Delta\rvert_{\text{пред.}} =2{,}5\ \sigma$.	$\lvert\Delta\rvert_{\text{пред.}} = 2\ \sigma$.
$P\ (\lvert\Delta\rvert < \sigma)$	0,68454	0,69128	0,71523
$P\ (\lvert\Delta\rvert < 2\sigma)$	0,95709	0,96651	1,00000
$P\ (\lvert\Delta\rvert < 2{,}5\sigma)$	0,99026	1,00000	
$P\ (\lvert\Delta\rvert < 3\sigma)$	1,00000		

Расхождение вероятностей усеченного нормального и нормального законов при $|\Delta|_{\text{пред.}} = 3\ \sigma$ мало, в третьей значащей цифре.

По сравнению с данными для допуска 3 σ при допуске 2,5 σ расхождение вероятностей усеченного нормального и нормального законов более существенно – во второй значащей цифре (до сотых).

Разности вероятностей усеченного нормального и нормального законов распределения представлены в табл. 3.

Таблица 3

Разности вероятностей

Интервал / $\|\Delta\|_{пред.}$	3σ	2,5 σ	2 σ
± σ	0,00185	0,00859	0,03254
± 2 σ	0,00255	0,00120	0,04255
± 2,5 σ	0,00268	0,01242	-
± 3 σ	0,00270	-	-

С увеличением «жесткости» допуска возрастает степень расхождения вероятностей усеченного нормального и нормального распределений. Так как теоретическое количество ошибок в заданном интервале определяется как произведение *n P*, (например, при вычислении критерия Пирсона χ^2), расхождение в три сотых на всем интервале $|\Delta| < \sigma$ при большом числе наблюдений может изменить теоретическое значение числа попадания ошибок в заданный интервал.

Возможно, в каких-то случаях использование вероятностей усеченного нормального закона позволит избежать повторных наблюдений.

Рассчитаем, как изменятся коэффициенты допусков, если предположить, что соответствующие им вероятности нормального закона будут приняты и для усеченного нормального распределения, $\Phi(t) = \Phi_1(t_1)$.

При $|\Delta|_{пред.} = 3\sigma$.

1. $P(|\Delta| < 2\sigma) = \Phi(2) = 0{,}95450$ – нормальный закон.

Для усеченного нормального закона

$\Phi_1(1{,}95) = 0{,}95139$,

$\Phi_1(2{,}00) = 0{,}95709$,

$t_1 = 1{,}98$. $t_1 < t$.

2. $P(|\Delta| < 2{,}5\sigma) = \Phi(2{,}5) = 0{,}98758$ – нормальный закон.

Для усеченного нормального закона

$\Phi_1(2{,}40) = 0{,}98627$,

$\Phi_1(2{,}45) = 0{,}98838$,

$t_1 = 2{,}43$. $t_1 < t$.

3. $P(|\Delta| < 3\sigma) = \Phi(3) = 0{,}99730$ – нормальный закон.

Для усеченного нормального закона

$\Phi_1(2{,}75) = 0{,}99673$,

$\Phi_1(2{,}80) = 0{,}99759$,

$t_1 = 2{,}78$. $t_1 < t$.

И во всех остальных случаях коэффициенты допусков усеченного нормального закона меньше соответствующих коэффициентов допусков нормального закона распределения. Области допустимых значений случайных ошибок измерений для нормального закона шире, чем для усеченного нормального распределения.

Вероятности усеченного логистического закона

Для сравнения результатов рассмотрим еще один закон распределения – логистический [7 - 14]. Ему подчиняются некоторые функции случайных ошибок геодезических измерений.

Функция логистического распределения

$$F(x) = \frac{1}{1+e^{-Z}}, \qquad (9)$$

где $z = (x-a)/\lambda$.

$M_x = a$ – математическое ожидание – параметр сдвига, $a \in (-\infty, \infty)$;

$\sigma_x = \pi\lambda / \sqrt{3}$ – среднее квадратическое отклонение,

λ – масштабный параметр, $\lambda \in (0, \infty)$.

Асимметрия $S_k = 0$, эксцесс $E \approx 1{,}2$.

Вероятность попадания СВ (X) в заданный интервал

$$P(x_1 \leq X < x_2) = F(x_2) - F(x_1). \qquad (10)$$

Вероятности логистического закона попадания случайных ошибок измерений в заданные интервалы приведены в табл. 4

Таблица 4

Вероятности логистического закона

Логистическое распределение
$P(\lvert\Delta\rvert < \sigma) = 0{,}7196$
$P(\lvert\Delta\rvert < 2\sigma) = 0{,}94822$
$P(\lvert\Delta\rvert < 2{,}5\sigma) = 0{,}97876$
$P(\lvert\Delta\rvert < 3\sigma) = 0{,}99138$

Для усеченного логистического закона

1.При $|\Delta|_{\text{пред.}} = 3\,\sigma$.

$a = -3\sigma$, $в = +3\sigma$, $M_x = 0$. $F(3\,\sigma) = 0{,}99569$, $F(-3\,\sigma) = 0{,}00431$, $c = 1{,}00869$.

2. При $|\Delta|_{\text{пред.}} = 2{,}5\,\sigma$.

a = - 2,5 σ, $в$ = +2,5 σ , M_x = 0. F(2,5 σ) = 0,98938, F(-2,5 σ) = 0,01062,
c = 1,02170.
3. При $|\Delta|_{пред.}$ = 2 σ
a = - 2 σ, в = +2σ , M_x = 0. F(2 σ) = 0,97411, F(-2 σ) = 0,02589, c = 1,05461.

Вероятности усеченного логистического закона приведены в табл.5.

Таблица5

Усеченное логистическое распределение

Вероятности/пределы	$\|\Delta\|_{пред.}$ = 3 σ.	$\|\Delta\|_{пред.}$ =2,5 σ.	$\|\Delta\|_{пред.}$ = 2 σ.
P (|Δ| < σ)	0,72590	0,73526	0,75894
P (|Δ| < 2σ)	0,95646	0,96880	1,00000
P (|Δ| < 2,5σ)	0,98727	1,00000	
P (|Δ| < 3σ)	1,00000		

Расхождение вероятностей усеченного логистического и логистического законов при $|\Delta|_{пред.}$ = 3σ имеет место в третьей значащей цифре, что является малой величиной.

При допуске 2,5σ расхождение вероятностей усеченного логистического и логистического законов более существенно – во второй значащей цифре (до сотых). Т.е., в целом такого же порядка, как вероятностей нормального и усеченного нормального распределения.

Максимальное расхождение вероятностей усеченного логистического и нормального законов составляет 0,043.

Заключение

1. Вероятности усеченного нормального и нормального законов отличаются незначительно. Максимальное расхождение составило 0,043.
2. Заменять вероятности допусков нормального закона на вероятности усеченного нормального закона не рекомендуется.
3. Можно использовать вероятности усеченного закона только в том случае, если это позволит избежать повторных полевых измерений.
4. Интервалы допустимых значений усеченного нормального закона меньше соответствующих интервалов нормального закона распределения.
5. Расхождения вероятностей усеченного логистического распределения и вероятностей логистического закона незначительны.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Большаков В.Д., Гайдаев П.А. Теория математической обработки геодезических измерений. – М.: Недра, 1977. – 367с.
2. Большаков В.Д., Маркузе Ю.И., Голубев В.В. Уравнивание геодезических построений. - М.: Недра, 1989. – 413 с.
3. Большев А.Н., Смирнов Н.Б. Таблицы математической статистики. – М.: Наука, 1965. – 464 с.

4. Видуев Н.Г., Кондра Г.В. Вероятностно – статистический анализ погрешностей измерений. – М.: Недра, 1969. – 320 стр.

5. Губарев В.В. Вероятностные модели: Справочник: В 2 ч.- Новосибирск: НЭТИ,1992. – 422 с.

6. Кендал М., Стьюарт А. Статистические выводы и связи. – М.: Наука, 1973. – 900 с.

7. Кокс Д., Хинкли Д. Теоретическая статистика. – М.: Мир, 1978. – 560 с.

8. Кремер Н. Ш. Теория вероятностей и математическая статистика. – М.: Юнити, 2007. – 550 с.

9. Маркузе Ю.И., Бойко Е.Г., Голубев В.В. Геодезия. Вычисление и уравнивание геодезических сетей. – М.: Картцентр – Геодезиздат, 1994. – 431с.

10. Статистический анализ независимых и зависимых случайных величин в геодезии: Отчет НИР (Заключительный)/ НГТУ; Руководитель Б.Ю. Лемешко.

Исполнители Б.Ю. Лемешко, Н.Б. Лесных, Г.И. Мизина, С.Н. Постовалов – № ГР 01.9.60 003120; Инв. № 02.9.80 001109. – Новосибирск, 1997. – 40 с.

11. Теория, методы и программное обеспечение задач статистического анализа независимых и зависимых случайных величин в геодезии. Отчет о НИР / НГТУ; Руководитель В.И. Денисов. Исполнители Б.Ю. Лемешко, Н.Б. Лесных, С.Н. Постовалов. № ТР 01.95. 000159; Инв. № 02.05. 0001199. – Новосибирск, 1994. – 40 с.

12. Лемешко Б.Ю. Статистический анализ одномерных наблюдений случайных величин. Программная система. – Новосибирск: НГТУ, 1995. – 125 с.

13. Лесных. Н.Б. Законы распределения случайных величин в геодезии: монография / Н.Б. Лесных. – Новосибирск: СГГА, 2005. – 128 с.

14. Лесных. Н.Б. Объекты статистического анализа в геодезии: монография / Н.Б. Лесных. – Новосибирск: СГГА. – 2010. – 128 с.