Краткие сообщения по физике № 5-1985 г.

АЛГОРИТМ ВЫЧИСЛЕНИЯ ВЕРХНЕГО ПРЕДЕЛА СЛУЧАЙНОЙ ВЕЛИЧИНЫ НА ЗАДАННОМ ДОВЕРИТЕЛЬНОМ УРОВНЕ

В. Ф. Грушин

УДК 519.25.9

Выводится общая формула для вычисления верхнего предела на 100 α-процентном доверительном уровне измеренной случайной величины $u = u_0 \pm \sigma$ при условии $u \geqslant 0$.

В физических исследованиях нередко возникает необходимость получать статистическую оценку сверху случайной величины, измеренной с большой погрешностью. В частности, для положительно определенных случайных величин, таких как сечение какого-либо процесса, ширина резонанса, выход редких событий и т. п., принято использовать в качестве подобной оценки значение верхнего предела на заданном (чаще всего 90-процентном) доверительном уровне /1,2/.

Ниже мы получим простой алгоритм вычисления такого верхнего предела для достаточно общего случая.

Пусть в результате измерения случайной величины u найдено:

$$u = u_0 \pm \sigma, \tag{1}$$

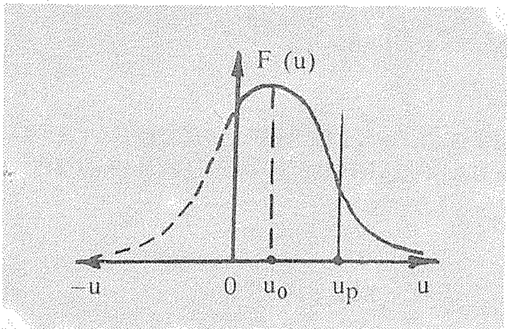
где u_0 означает оценку среднего значения, а σ – среднеквадратичная ошибка. При этом предполагается, что u имеет нормальное распределение F(u) (например, гауссово). Вопрос об оценке верхнего предела измеренной величины u встает при $u_0 \lesssim \sigma$.

На рис. 1 изображена ситуация, соответствующая результату (1) для положительно определенной величины u* . Функция распределения F(u) есть обычный гауссиан:

$$F(u) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left[\frac{-(u-u_0)^2}{2\sigma^2}\right].$$

---

* Если из-за большой погрешности измерений обнаружено отрицательное u_0, то в этом случае результат обычно интерпретируют как $0 \pm \sigma$.



Р и с. 1.

Задача состоит в том, чтобы найти такое значение u_p, которое является верхним пределом величины u (1) на 100 α-процентном доверительном уровне. Исходным соотношением для поиска u_p по определению является

$$\int_{u_p}^{\infty} F(u)\,du \Big/ \int_{0}^{\infty} F(u)\,du = 1 - \alpha. \tag{2}$$

Выразим числитель и знаменатель через стандартизованный гауссиан $f(\tilde{u}) = (2\pi)^{-1/2} e^{-\tilde{u}^2/2}$ (с нулевым средним и единичной дисперсией), для которого имеются различные численные таблицы в руководствах по математической статистике. Получаем

$$\int_{u_p}^{\infty} F(u)\,du = \int_{(u_p - u_0)/\sigma}^{\infty} f(\tilde{u})\,d\tilde{u},$$

$$\int_{0}^{\infty} F(u)\,du = \int_{-\infty}^{\infty} f(\tilde{u})\,d\tilde{u} - \int_{u_0/\sigma}^{\infty} f(\tilde{u})\,d\tilde{u} = 1 - \int_{u_0/\sigma}^{\infty} f(\tilde{u})\,d\tilde{u}.$$

Теперь вместо выражения (3) имеем

$$\int_{(u_p - u_0)/\sigma}^{\infty} f(\tilde{u})\,d\tilde{u} \Big/ \left[1 - \int_{u_0/\sigma}^{\infty} f(\tilde{u})\,d\tilde{u}\right] = 1 - \alpha. \tag{3}$$

Интеграл $\int_a^{\infty} f(\tilde{u})\,d\tilde{u}$ – площадь под стандартизованным гауссианом, отсекаемая вертикалью, проходящей через координату u = a. Обычно в численных таблицах нормального распределения принято приводить значения $p(a) = 2\int_a^{\infty} f(\tilde{u})\,d\tilde{u}$, которые соответствуют оценке вероятности в пределах двухстороннего доверительного интервала $(-a;\ a)$. Поэтому для использования таких таблиц (см., например, /3/) в наших целях обратим внимание на то, что $\int_a^{\infty} f(\tilde{u})\,d\tilde{u} = p(a)/2$. На этом основании перепишем (3) в виде:

$$(1/2)\,p((u_p - u_0)/\sigma)/[1 - 1/2p(u_0/\sigma)] = 1 - \alpha.$$

Отсюда получаем

$$P\left[\frac{u_p - u_0}{\sigma}\right] = 2(1 - \alpha)\left[1 - \frac{1}{2}P\left(\frac{u_0}{\sigma}\right)\right]. \qquad (4)$$

Выражение (4) является конечным алгоритмом для определения u_p по заданным u_0, σ и α* .

В качестве примера определим u_p для частного случая: $u_0/\sigma = 1$; $\alpha = 0{,}9$. Из таблиц находим $p(1) = 0.32$, что приводит (4) к числовому равенству $p((u_p - u_0)/\sigma) = 0{,}168$. Вновь из тех же таблиц находим $(u_p - u_0)/\sigma \cong 1{,}4$ и окончательно $u_p \cong 2{,}4\ \sigma$, что является хорошо известным результатом, используемым на практике.

Рассмотрим еще один достаточно популярный случай: $u_0/\sigma \ll 1$; $\alpha = 0{,}9$. Здесь $p(u_0/\sigma) = 1$, и равенство (4) приобретает вид $p((u_p - u_0)/\sigma) = 0{,}10$.

С помощью таблиц находим $(u_p - u_0)/\sigma = 1{,}61$, т. е. $u_p = 1{,}61\ \sigma$, что находится в полном согласии с известным и в этом случае результатом.

Очевидно, что вычисление искомой величины u_p посредством алгоритма (4) легко осуществить и с помощью ЭВМ.

Поступила в редакцию 25 января 1985 г.

ЛИТЕРАТУРА

1. Aubert J. J. et al. Phys. Lett. **106B**, 419 (1981).
2. Bollini D. et al. Nucl. Phys., **B199**, 27 (1982).
3. Худсон Д. Статистика для физиков. М., Мир. 1967.
4. Helene O. Nucl. Instr. Meth., 228, 120 (1984).

* Когда $F(u)$ – дискретное распределение Пуассона, задачу о верхнем пределе не удается решить аналитически: так, в работе /4/ величина u_p рассчитана и табулирована для ряда конкретных случаев.