

РОССИЙСКАЯ АКАДЕМИЯ НАУК

СИБИРСКОЕ ОТДЕЛЕНИЕ

А В Т О М Е Т Р И Я

2006, том 42, № 6

УДК 519.25 : 577.15/17:681.142.1

**НЕПАРАМЕТРИЧЕСКИЙ МЕТОД
ПОСТРОЕНИЯ ЗАВИСИМОСТИ «ДОЗА–ЭФФЕКТ»**

А. Л. Осипов, С. Н. Аношкин

Институт прикладной информатики

Новосибирского государственного университета экономики и управления, г. Новосибирск

E-mail: osipov@nsaem.ru

Разработана компьютерная система, моделирующая функцию эффективности по результатам независимых единичных испытаний на основе непараметрического метода ядерной оценки регрессии.

Введение. На современном этапе в токсикометрии используются величины доз, которые вызывают появление эффекта, учитываемого в экспериментальной группе тест-объектов с заданной вероятностью 0,05; 0,16; 0,5; 0,84; 0,95. Такие дозы получили название эффективных доз: ED_5 , ED_{16} , ED_{50} , ED_{84} , ED_{95} . Общая зависимость вероятности появления эффекта от воздействия заданной дозы определяется как функция эффективности. В качестве единого показателя для сравнения тестируемых препаратов принята среднеэффективная доза ED_{50} , определяемая из функции эффективности. Проблема вероятностной оценки токсического и других эффектов рассматривается как одна из важнейших в токсикометрии и других разделах биологии и широко разрабатывается исходя из задач экспериментальной практики [1–3]. Открытие Гауссом закона нормального распределения послужило основой для разработки пробит-анализа, который в разных модификациях [4] используется в качестве основного метода определения среднеэффективных доз. В токсикометрии важное значение отводится методам определения среднеэффективных доз, так как они являются теми решающими факторами, от которых зависит способ планирования экспериментов, порядок формирования и объем исходных данных, а в конечном итоге качество, эффективность и достоверность искомых показателей токсичности. По этим признакам проблему токсикометрической оценки показателей токсичности можно рассматривать как фундаментальную проблему теоретической токсикологии, имеющей прикладное значение для других разделов биологии и медицины.

Анализ существующих подходов к определению среднеэффективной дозы. Известно несколько экспериментальных экспресс-методов определения средних летальных доз или концентраций. Эти методы различаются количеством используемых тест-объектов и схемой проведения эксперимен-

тов. В работе [5] была выполнена экспериментальная оценка точности определения ED_{50} с помощью различных экспресс-методов (пробит-анализа, методов Прозоровского, Дейхмана и Ле Бланка, Фрумина, сплайн-интерполяции и др.). Анализируя существующие методы расчета среднеэффективной дозы, отметим следующее.

1. Все существующие методы расчета среднеэффективной дозы используют в качестве исходных данных значения частот эффектов, наблюдаемых в однородных группах тест-объектов после испытания на них заданных доз исследуемых веществ. Это предполагает предварительное планирование токсикологического эксперимента для формирования исходных данных.

2. Существующие методы расчета среднеэффективной дозы строятся на гипотезе о нормальном распределении функции эффективности, поэтому в этих методах решается частная задача оценки вида распределения функции эффективности. Основываясь на методологическом подходе, предложенном в [6], построим математическую модель для компьютерного вычисления функции эффективности.

Математическая модель построения функции эффективности.

Пусть X обозначает случайную величину или минимальную дозу, при воздействии которой у тест-объекта с полной вероятностью проявится положительный эффект, например летальный исход. Назовем эту дозу минимальной абсолютно эффективной (летальной) дозой. Обозначим через Y экспериментально испытанную дозу, а через W наблюдавшийся у тест-объекта эффект. С точки зрения теории вероятностей результаты токсикологического эксперимента состоят в том, что после испытания заданной дозы Y у тест-объекта появится либо положительный ($W = 1$), либо отрицательный ($W = 0$) эффект. Если испытанная доза больше гипотетической $Y > X$, то регистрируется положительный эффект, и наоборот. Таким образом, показатель эффекта W служит индикатором события $\{Y > X\}$.

Биологическая суть вышеизложенного становится более понятной из следующих пояснений. Например, для каждого яда теоретически существует минимальная доза, которая вызывает у конкретного тест-объекта летальный эффект. Экспериментально определить эту дозу для каждого животного невозможно. Если животное в эксперименте выжило, то оно получило дозу яда, заведомо меньшую минимальной абсолютно летальной дозы. Для каждого животного эта доза будет различной, что определяется индивидуальной чувствительностью особей биологического вида к тестируемому препарату. С точки зрения теории вероятностей минимальная абсолютно эффективная доза представляет собой случайную величину с функцией распределения $F(x) = P(X < x)$, которую требуется найти по выборке $(Y_1, W_1), (Y_2, W_2), \dots, (Y_n, W_n)$. Если выразить через α вероятность того, что заданная доза X_α будет являться гипотетической минимальной дозой, при действии которой у тест-объектов проявятся положительные эффекты, то из уравнения $F(X_\alpha) = \alpha$ можно определить значение X_α эффективной дозы, вызывающей у тест-объекта заданную вероятность появления эффекта. Значения найденных доз в этом случае будут представлять собой категории эффективных доз: $ED_5, ED_{16}, ED_{50}, ED_{84}, ED_{95}$ и т. д. Функция $F(x)$ и является функцией эффективности исследуемого вещества.

Рассмотрим условное математическое ожидание величины W , когда значение введенной дозы Y фиксировано, т. е. $Y = y$. При фиксированном значении Y условное математическое ожидание W согласно теории вероятностей составляет $E(W | Y = y) = P(W = 1 | Y = y)$. Последняя вероятность соответст-

вует $P(W = 1 | Y = y) = P(X < y | Y = y)$. Если величины X и Y независимы, то условная вероятность $P(X < y | Y = y)$ равна безусловной вероятности $P(X < y) = F(y)$. Таким образом, $E(W | Y = y) = F(y)$. Иными словами, можно оценивать не функцию распределения, а условное математическое ожидание $E(W | Y = y)$.

Суть метода построения функции эффективности заключается в определении вероятности эффекта в середине заданного диапазона доз при помощи ядерной оценки регрессии. Если в заданном диапазоне находится n значений доз Y_i из общей выборки с соответствующими индикаторами альтернативного эффекта W_i , то оценка вероятности эффекта в середине диапазона может быть получена по формуле

$$F_h(x) = \frac{\sum_{i=1}^n K\left(\frac{Y_i - x}{h}\right) W_i}{\sum_{i=1}^n K\left(\frac{Y_i - x}{h}\right)},$$

где константа h задает величину окна просмотра, а функция

$$K(x) = \frac{3(1-x^2)}{4}, \quad -1 \leq x \leq 1,$$

называется ядром Епанечникова [7]. Кроме ядра Епанечникова используются и другие ядра, эффективность которых продемонстрирована в [8]. В этой же работе показана оптимальность ядра Епанечникова. При увеличении объема выборки оценка $F_h(x)$ приближается к нормальному распределению с дисперсией

$$\sigma^2 = \frac{1}{\sqrt{nh}} \frac{3F(x)(1-F(x))}{5f(x)}$$

($f(x)$ – плотность распределения функции эффективности $F(x)$). Для оценки дисперсии нужно в формулу для σ^2 подставить вместо $F(x)$ функцию $F_h(x)$, а вместо $f(x)$ ее оценку в точке x :

$$f_h(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{Y - x_i}{h}\right).$$

Если ядерная оценка вычисляется только в точках наблюдений (Y_i, W_i) , $i = 1, 2, \dots, n$, тогда при $h \rightarrow 0$

$$F_h(x) \rightarrow \frac{K(0)W_i}{K(0)} = W_i,$$

следовательно, малая ширина окна воспроизводит данные. При $h \rightarrow \infty$

$$K\left(\frac{Y_i - x}{h}\right) \rightarrow K(0).$$

Тогда

$$F_h(x) \rightarrow \sum_{i=1}^n K(0)W_i / \sum_{i=1}^n K(0) = \frac{1}{n} \sum_{i=1}^n W_i.$$

Слишком большие значения ширины окна приводят к чрезмерному сглаживанию кривой эффективности – среднему арифметическому значений переменной отклика. Возникает оптимизационная задача выбора параметра сглаживания. В [9] приводятся результаты исследования сходимости ядерных оценок. В частности, отмечено, что оптимальная скорость сходимости ядерной оценки плотности обеспечивает выбор параметра h на уровне $h = (\alpha/(\beta n))^{1/5}$, где

$$\alpha = \int K^2(t) dt, \quad \beta = \left(\int t^2 K(t) dt \right)^2 \int (f''(t))^2 dt.$$

При этом также предполагается, что $h \rightarrow 0$, $nh \rightarrow \infty$ при $n \rightarrow \infty$; $f(t)$ – ограниченная плотность, имеющая две непрерывные производные; $\int (f''(t))^2 dt < \infty$.

Данный результат можно использовать для определения оптимального параметра h . В приведенной выше формуле неизвестным параметром является плотность $f(t)$, используемая для определения коэффициента β . Заменим данную плотность ее ядерной оценкой $f_h(x)$. Поскольку ядерная оценка зависит от параметра h , то для поиска оптимального значения параметра необходимо организовать итеративную процедуру проведения расчетов.

На первом этапе выбираем произвольное значение $h_1 < 1$. Для этого значения строим ядерную оценку плотности $f_{h_1}(x)$ и на ее основании рассчитываем значения коэффициентов α_1, β_1 . Далее подставляем полученные значения коэффициентов в формулу $h = (\alpha/(\beta n))^{1/5}$ и вычисляем новое значение параметра $h_2 = (\alpha_1/(\beta_1 n))^{1/5}$. На основании этого значения параметра строим новую ядерную оценку $f_{h_2}(x)$. Полученное значение плотности используем вновь для расчета коэффициентов α, β . Повторяем данную процедуру до тех

Название химического вещества	ED ₅₀		
	эксперимент	пробит-анализ	непараметрический метод
Ацетилхолин	4,25 ± 0,29	4,554391	4,561538
Адреналин	3,68 ± 0,38	3,500471	3,599750
Фенамин	86,00 ± 8,70	57,181177	85,505580

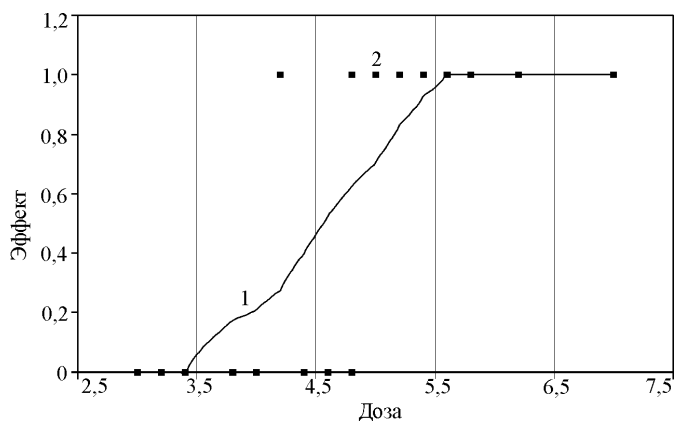


Рис. 1

пор, пока не будет выполняться условие сходимости результата вычисления оптимизируемого параметра h , а именно

$$|h_m - h_{m-1}| < \varepsilon,$$

где ε – малое число, определяющее заданную точность проведения расчетов. Эта процедура встроена как программный модуль в компьютерную систему для моделирования зависимости «доза–эффект».

Результаты компьютерного вычисления функции эффективности с помощью непараметрического подхода (см. таблицу). На рис. 1, 2 приведены вычисленные с помощью компьютерной системы графики функций эффективности для следующих химических веществ.

График функции эффективности для ацетилхолина (см. рис. 1, кривая 1 – прогноз, 2 – исходные данные) близок к функции нормального распределения, поэтому результаты вычислений эффективных доз пробит-анализом и непараметрическим методом совпадают.

Совсем по-другому обстоят дела с фенамином (см. рис. 2, кривая 1 – оценка, 2 – исходные данные). Функция эффективности для этого вещества существенно отличается от функции нормального распределения, поэтому получаются значительные различия при вычислении эффективных доз методами пробит-анализа и непараметрическим.

В сравнении с классическими методами пробит-анализа непараметрический метод дает лучшие результаты при вычислении функции эффективности.

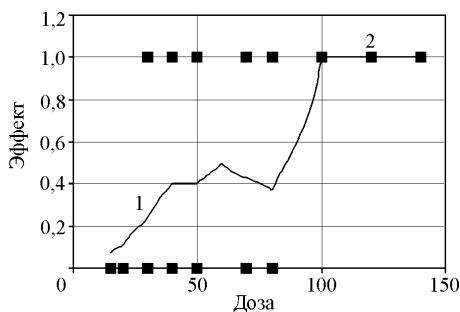


Рис. 2

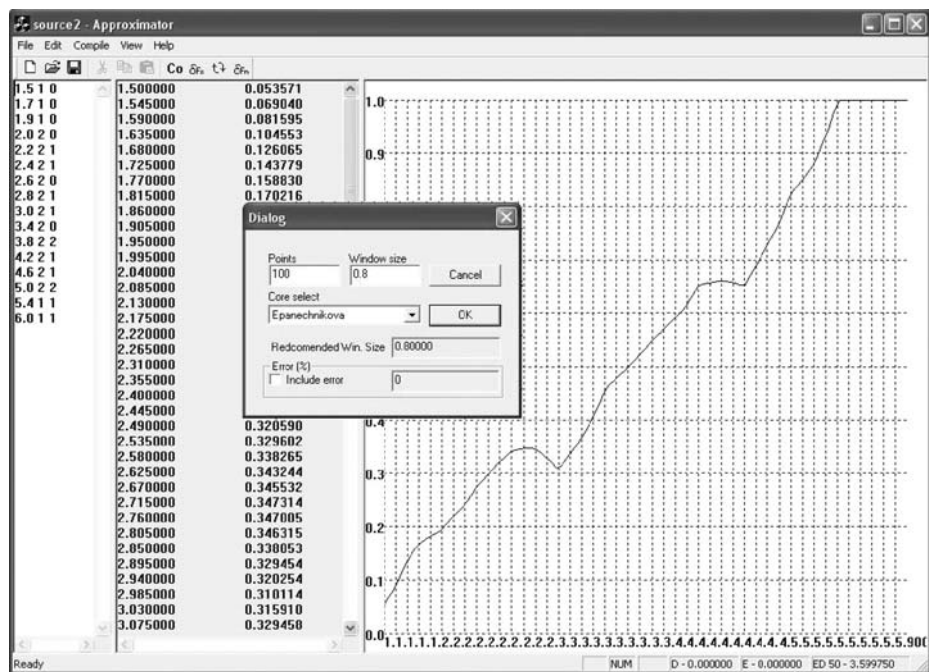


Рис. 3

На рис. 3 представлен интерфейс программной системы для вычисления функции эффективности.

Заключение. Разработана компьютерная система, позволяющая строить функцию эффективности для расчета среднеэффективных доз на основе непараметрического подхода, а также методов пробит-анализа. В сравнении с методами пробит-анализа непараметрический метод показал более высокую эффективность в расчетах категорий эффективных доз. Для веществ, функция эффективности которых близка к функции нормального распределения, компьютерная система дает результаты, близкие к полученным классическим методом пробит-анализа, а при значительных различиях функций дает результаты, более близкие к экспериментальным, чем классический метод пробит-анализа.

СПИСОК ЛИТЕРАТУРЫ

1. Новиков С. М., Левченко Н. И., Мельникова Н. Н., Фурсова Т. Н. Определение вероятностных значений пороговых уровней воздействия вредных веществ // Гигиена и санитария. 1989. № 9. С. 46.
2. Копанев В. А., Гинзбург Э. Х., Семенова В. Н. Метод вероятностной оценки токсического эффекта. Новосибирск: Наука, 1988.
3. Прозоровский В. Б., Фрумин Г. Т. Аналитический вариант экспрессного метода одной точки для определения средних эффективных доз // Экспериментальная и клиническая фармакология. 1992. 55, № 6. С. 61.

4. **Russel R. W., Robertson J. L., Savin N. E.** A computer program for multiple probit or logit analysis // Bull. Entomol. Soc. Am. 1981. **27**, N 3. P. 210.
5. **Фрумин Г. Т., Загребин Ф. О., Тонкопий В. Д.** Точность экспресс-методов определения средних летальных концентраций // Хим.-фарм. журн. 1993. **27**, № 6. С. 17.
6. **Криштопенко С. В., Тихов М. С., Попова Е. Б.** Парадоксальная токсичность. Нижний Новгород: Изд-во Нижегородской гос. мед. акад., 2001.
7. **Епанечников В. А.** Непараметрическая оценка многомерной плотности вероятности // Теория вероятностей и ее применения. 1969. **14**, № 1. С. 156.
8. **Хардле В.** Прикладная непараметрическая регрессия: Пер. с англ. М.: Мир, 1993.
9. **Деврой Л., Дьерфи Л.** Непараметрическое оценивание плотности. L1-подход: Пер. с англ. М.: Мир, 1988.

Поступила в редакцию 18 января 2006 г.
