§3. Интерполяция данных

Практически всегда выборки случайных чисел (полученные в результате эксперимента или сгенерированные в рамках методов Монте-Карло) хранятся на компьютерах в виде массивов, т.е. дискретных данных $y_i(x_i)$. При решении различного рода задач имеется необходимость оперировать непрерывной случайной функцией y(x) одной или нескольких переменных, что соответствует непрерывному случайному процессу или случайному полю. Для получения y(x) необходимо построить интерполяционную зависимость в промежутках между точками.

Интерполяция-экстраполяция, напомним, подразумевает «соединение» точек выборки данных (x_i,y_i) кривой той или иной степени гладкости. По определению интерполяция означает построение функции f(x), аппроксимирующей зависимость y(x) в промежуточных точках (между x_i). Поэтому интерполяцию еще по-другому называют *аппроксимацией*. В точках x_i значения интерполяционной функции должны совпадать с исходными данными, т.е. $f(x_i)=y(x_i)$.

Линейная интерполяция

Самый простой вид интерполяции – кусочно-постоянная. Ее смысл состоит в том, что на каждом промежутке между экспериментальными точками f(x) представляет собой константу, равную значению случайной величины на левой (рис. 26) или (реже) правой границе.

Минус такого подхода заключается в том, что ни полученная интерполяция, ни ее производная не являются непрерывными функциями. Между тем, в нашем случае данных о курсе валют такая интерполяция является наиболее правильной, т.к. отвечает специфике экспериментальных данных. Ведь официальный курс устанавливается Центробанком в некоторый момент времени и действует до его следующего объявления – на следующий день или через несколько дней (если текущая дата выпадает на праздники).

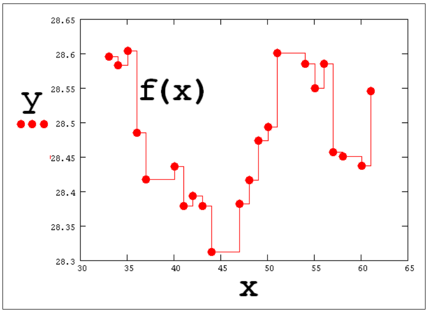

Рис. 26. *Кусочно-постоянная интерполяция*

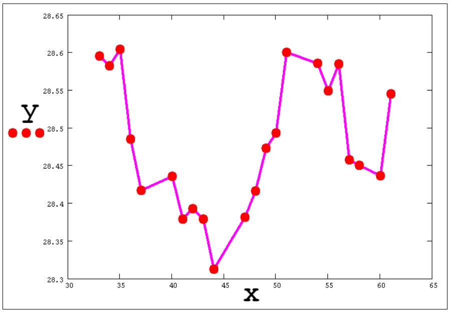

Рис. 27. *Кусочно-линейная интерполяция*

Чуть сложнее выглядит кусочно-линейная интерполяция, которая представляет искомую зависимость в виде ломаной линии. Интерполирующая функция f(x) состоит из отрезков прямых, соединяющих данные «от точки к точке» (рис. 27).

Приведем формулу кусочно-линейной интерполяции:

$$f(x)=y_i\cdot\frac{x-x_{i+1}}{x_i-x_{i+1}}+y_{i+1}\cdot\frac{x-x_i}{x_{i+1}-x_i} \qquad (7)$$

и оценку ее погрешности, т.е. максимального отклонения f(x) от неизвестной «настоящей» функции F(x):

$$|f(x)-F(x)|\leqslant\frac{h^2}{8}\cdot\max(F''(x)) \quad . \qquad (8)$$

Таким образом, погрешность определяется второй производной функции F(x) и убывает пропорционально квадрату шага h между данными.

Сплайн-интерполяция

В большинстве практических приложений желательно соединить экспериментальные точки не ломаной линией, а гладкой кривой. Лучше всего для этих целей подходит интерполяция *сплайнами*, т. е. фрагментами полиномов. Смысл сплайн-интерполяции заключается в том, что в каждом промежутке между узловыми точками (на каждом шаге интерполяции) осуществляется аппроксимация в виде определенной полиномиальной зависимости f(x). При этом для каждого шага получается свой полином, причем его коэффициенты подбираются такими, чтобы на границах шага выполнялись условия сшивки. А именно, если применяются сплайны в виде полиномов степени m, то несложно показать, что их коэффициенты можно выбрать так, чтобы обеспечить непрерывность производных порядка до (m-1)-й включительно.

Наиболее часто применяются кубические сплайны, т.е. полиномы третьей степени (кубические параболы):

$$f(x)=ax^3+bx^2+cx+d. \tag{9}$$

Коэффициенты a,b,c,d рассчитываются независимо для каждого промежутка интерполирования, исходя из значений y_i в соседних точках. Участки парабол называются *кубическими сплайнами*.

Приведем формулы для расчета кубической сплайн-интерполяции. Искомая функция (9) на промежутке между x_i и x_{i+1} вычисляется следующим образом:

$$\begin{aligned} f(x)= & \frac{(x_{i+1}-x)^2\cdot\left(2(x_i-x)+h\right)\cdot y_i}{h^3}+\frac{(x_{i+1}-x)^2\cdot(x-x_i)\cdot m_i}{h^2} \\ & +\frac{(x_{i+1}-x)^2\cdot\left(2(x_i-x)+h\right)\cdot y_{i+1}}{h^3}+\frac{(x_{i+1}-x)\cdot(x-x_i)^2\cdot m_{i+1}}{h^2} \end{aligned}, \tag{10}$$

где

$$m_i=\frac{(y_{i+1}-y_{i-1})}{2h}, \tag{11}$$

$$m_0=\frac{(-3y_0+4y_1-y_2)}{2h}, \tag{12}$$

$$m_N = \frac{(3y_N - 4y_{N-1} - y_{N-2})}{2h} \quad . \tag{13}$$

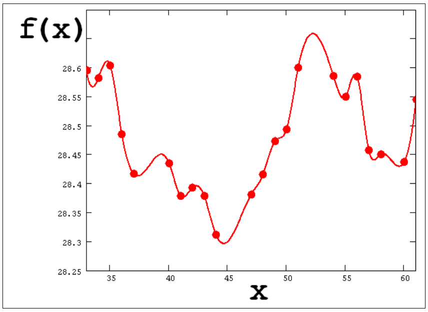

Рис. 28 *Сплайн-интерполяция*

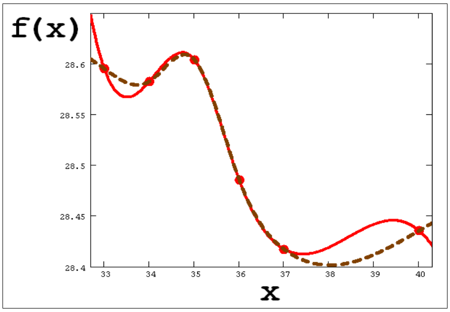

Рис. 29. *Сплайн-интерполяция (два варианта)*

Кубическая сплайн-интерполяция обеспечивает равенство в узлах не только самих соседних параболических интерполирующих функций (сплайнов), но и их 1-х и 2-х производных. Благодаря этому сплайн-интерполяция выглядит очень гладкой (рис. 28). Отметим, что формулы (10-13) – это не единственный способ построения кубических сплайнов (т.е. сшитых фрагментов кубических парабол, обладающих непрерывной 2-й производной). Имеется некоторый произвол в определении сплайн-интерполяции, который в подавляющем большинстве задач не приводит к сколь нибудь значимому эффекту (на рис. 29 изображен фрагмент графика, который интерполируется двумя способами).