УДК 621.396 + 519.2

КРИТЕРИЙ СОГЛАСИЯ НА ЭКСТРЕМАЛЬНЫХ СТАТИСТИКАХ

А.Ф. Барышников, аспирант МТУСИ
В.Н. Репинский, декан факультета «Информационные технологии» МТУСИ, к.т.н.

Построение прогноза поведения большой системы при большом числе случайных воздействий может быть сведено к прогнозированию параметров трендов временных рядов, представляющих собой выборочные значения случайного процесса с неизвестными параметрами. Как было указано в [2], такой прогноз основывается на гипотезе о непринадлежности двух скользящих выборок одному распределению генеральной совокупности. Основным инструментом для построения прогноза является выбор подходящей статистики и построение системы принятия решений.

В статье рассмотрен метод прогнозирования, позволяющий предсказать изменение тренда сигнала. Такой прогноз может быть полезен для коротковолновых (ионосферных) систем, когда их работа затруднена помехами различного происхождения.

Согласно теореме Гнеденко, существуют только три типа асимптотических распределений наибольшего значения выборки $X_{(n)}$ при бесконечном увеличении n (объема выборки). Из нее также следует, что наибольший $X_{(n)}$ и наименьший $X_{(1)}$ элементы выборки (экстремальные значения) асимптотически взаимно независимы.

Второй и третий типы существуют для исходных распределений, у которых случайная переменная имеет верхнюю или нижнюю границу. Очевидно, что эти значения являются границами и для экстремальных значений случайного процесса. Граница входит в асимптотические распределения как параметр k, зависящий от объема выборки. Примером распределения с границей может служить равномерное распределение, имеющее ограничение как сверху, так и снизу. Второй и третий типы распределений подробно рассмотрены в [2].

Рассмотрим первый тип распределения — самый важный из трех. Ему подчиняются оба экстремальных значения нормального и логистического распределений и наибольшие значения экспоненциального, логарифмически нормального и гамма-распределения. Точные формулы распределения экстремальных значений крайне неудобны для расчетов [1, с. 69].

Поэтому используют асимптотические распределения, применимые для достаточно больших объемов выборок n. Асимптотические распределения экстремальных значений имеют простой аналитический вид, поскольку они зависят в явном виде от таких характеристик случайного процесса, как экстремальные значения и интенсивности. Эти параметры, в свою очередь, зависят от асимптотических свойств краев исходной функции распределения.

Характеристические наибольшие и наименьшие значения u_n и u_1 выражаются как квантили исходного распределения:

$$F(u_1)=\frac{1}{n};\ F(u_n)=1-\frac{1}{n}.$$

Экстремальные интенсивности:

$$\alpha_1=nf(u_1);\ \alpha_n=nf(u_n).$$

Асимптотические функции плотности и распределения наибольшего значения (первый тип):

$$F_n(x)=\exp\left[-\exp[-\alpha_n(x-u_n)]\right],\ -\infty<x<\infty;$$

$$f_n(x)=\alpha_n\exp[-\alpha_n(x-u_n)]F_n(x).$$

Асимптотические функции плотности и распределения наименьшего значения (первый тип):

$$F_1(x)=\exp\left[-\exp[\alpha_1(x-u_1)]\right],\ -\infty<x<\infty;$$

 ISSN 0013-5771. «ЭЛЕКТРОСВЯЗЬ», № 1, 2009

$$f_1(x) = \alpha_1 \exp[\alpha_1(x - u_1)] F_1(x).$$

Экстремальное отношение — это отношение максимального значения выборки к минимальному:

$$q_n = -\frac{x_{(n)}}{x_{(1)}}.$$

Выражение для функции плотности экстремального отношения можно найти, зная функции плотности минимального и максимального значений:

$$f_q(q) = \int_{-\infty}^{\infty} |x_1| f_n(-qx_1) f_1(x_1) dx_1.$$

Введем новую статистику — отношение экстремальных отношений

$$Q = q_2 / q_1,$$

которая определяется по двум выборкам. Оказалось [2], что она очень чувствительна к изменению характера случайного процесса.

Функции плотности и распределения отношения отношений Q:

$$f(Q) = \int_{-\infty}^{\infty} |q| f_q(Qq) f_q(q) dq; \; F(Q) = \int_{-\infty}^{Q} f(Q') dQ'.$$

Помимо аргумента Q, у этих функций есть четыре параметра (a_1, a_n, u_1, u_n), которые зависят от объема выборки n и от функции распределения случайного процесса.

Рассмотрим частный случай для первого типа асимптотических распределений, когда исходное распределение **нормальное**. Для такого случая была найдена аппроксимирующая функция распределения отношения экстремальных отношений, показавшая достаточно высокую точность аппроксимации:

$$f_{appr}(Q) = \frac{\Phi Q^{\Phi-1}}{(1+Q^{\Phi})^2}, \; Q \geq 0;$$

$$F_{appr}(Q) = \frac{Q^{\Phi}}{1+Q^{\Phi}}, \; Q \geq 0.$$

Для каждого n будет свое оптимальное значение Ф. В качестве критерия оптимальности возьмем минимальную сумму квадратов разностей:

$$\sigma(\Phi) = \frac{1}{i_{\max}} \sum_{i=1}^{i_{\max}} \left[f_{appr}(Q_i, \Phi) - f(Q_i) \right]^2,$$

где $i_{\max}$ — количество точек, по которым будет производиться аппроксимация;

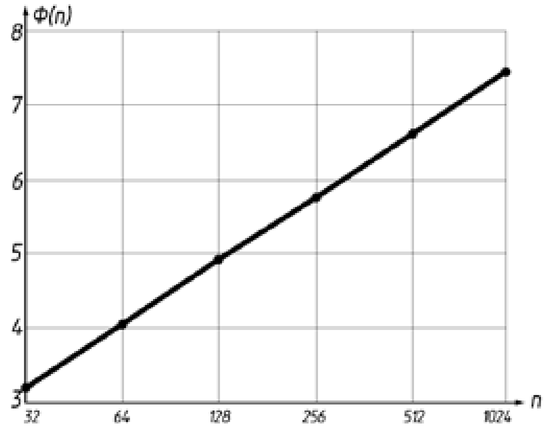

Рис. 1

$$\Phi_{opt} = \min \sigma(\Phi).$$

Решив уравнение $\frac{d\sigma(\Phi)}{d\Phi} = 0$, построим график $\Phi_{opt}(n)$ в логарифмическом масштабе (рис. 1).

Кривая отлично аппроксимируется линейной функцией вида:

$$\Phi(n) = C_0 + C_1 \ln(n).$$

Коэффициенты C_0 и C_1 найдем, используя тот же критерий оптимальности. Окончательный вид $\Phi(n)$:

$$\Phi(n) = -1{,}063 + 1{,}231 \ln(n).$$

Погрешность аппроксимации мала и вблизи рабочих значений, т. е. при $Q \approx 1$ стремится к 0 (рис. 2).

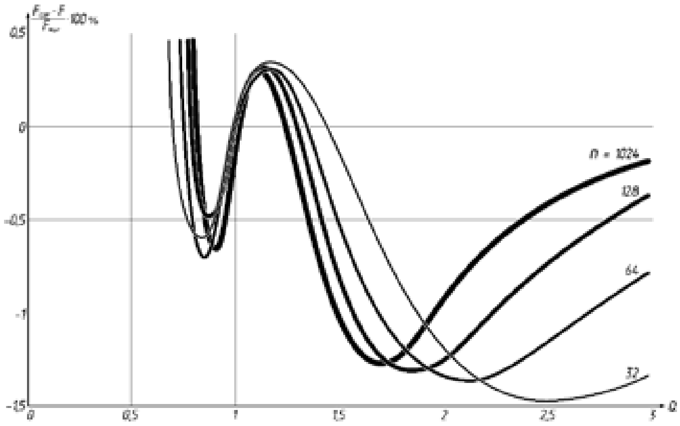

Рис. 2

Аналитические выражения для второго и третьего типов асимптотических распределений были распространены на случай любого значения k и сильно упрощены.

В [2] было найдено решение для $k = 2^m$:

$$F_k(Q) = 1 - \frac{1 + kQ^k \ln(Q) - Q^k}{(Q-1)^2 \prod_{i=0}^{\log_2 k - 1} \left(1 + Q^{2^i}\right)^2}.$$

Данное выражение громоздко и применимо только для ряда значений k. Существуют также трудности при нахождении предела функции в точке $Q = 1$.

Выражение для всех значений k:

$$F_k(Q) = Q^{-k} \left(\frac{\ln(Q^{-k})}{(Q^{-k} - 1)^2} + \frac{Q^{2k}}{Q^k - 1} \right) = \frac{A(A - 1 - \ln A)}{(A-1)^2},$$

где $A = Q^k$, а $\lim_{Q \to 1} F_k(Q) = 0{,}5$.

Полученное выражение — основа для применения критерия экстремальных отношений при прогнозировании изменений тренда случайного процесса.

ЛИТЕРАТУРА

1. **Боярский А.Я.** Введение в теорию порядковых статистик. — М.: Статистика, 1970. — 412 с.
2. **Репинская Т.В.** Непараметрический критерий согласия на экстремальных статистиках // Электросвязь. — 2004. — №8. — С.16—17.

Получено 04.02.08