УДК 519.2–621.396

Г. Ю. СИДОРЕНКО, канд. техн. наук, доц. НТУ «ХПИ;
А. С. МАЗМАНИШВИЛИ, д-р физ.-мат. наук, профессор, профессор СГУ, Сумы

РАСПРЕДЕЛЕНИЕ СТЬЮДЕНТА В СЛУЧАЕ КОРРЕЛИРОВАННЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ

В статті розглянуто типові задачі математичної статистики, що відносяться до корельованих послідовностей (часовим відлікам) з заданим фіксованим рівнем статистичного зв’язку між ними. Одержані аналітичні вирази, що описують розподіл Стьюдента у випадку корельованих відліків. Представлені графіки функції та розподілу Стьюдента при наявності кореляції.

Рассмотрены типичные задачи математической статистики, отнесенные к коррелированным последовательностям (временным отсчетам) с заданным фиксированным уровнем статистической связи между ними. Получены аналитические выражения, описывающие распределение Стьюдента для случая коррелированных отсчетов. Представлены графики функции и распределения Стьюдента при наличии корреляции.

The typical problems of mathematical statistics, related to correlated sequences (to time counts) with a given fixed level of statistical correlation between them, are considered. The analytic expressions featuring distributions Student's ones in the case of correlated counts are obtained. The charts of functions and distributions Student's in the case of correlated counts are introduction.

Введение. Многие типичные задачи математической статистики рассматриваются, в частности, в предположении о независимости фигурирующих в них элементов последовательностей (отсчетов) $x_1,\ldots,x_n$. К таким задачам, прежде всего, относятся распределение χ^2, распределение Стьюдента [1-3]. Между тем, вполне возможна ситуация, когда наблюдаемые данные или данные, предназначенные для статистической обработки, в той или иной степени зависимы, что часто встречается в различных экспериментах. Возможно также, что информация о коррелированности отсчетов отсутствует или неизвестна к моменту обработки результатов эксперимента. В этой связи возникает вопрос о степени влияния корреляции данных на распределения рассматриваемых в задачах математической статистики величин (прямые задачи).

Целью настоящей работы является, таким образом, изучение степени влияния коррелированности отсчетов на статистический закон распределения Стьюдента, при этом в качестве источника отсчетов будет использована коррелированная последовательность.

Основная часть. Рассмотрение будет проведено при следующих предположениях:

а) Используемые ниже наборы отсчетов $\{x_1,\ldots,x_n\}$ – совокупность сечений объемом n, взятых с постоянным временным шагом τ из реализации нормального Марковского процесса Орнштейна-Уленбека (ОУ-процесса) [4–6] $x(t)$ с интенсивностью σ и декрементом ν;

б) Нулевой отсчет y_0 случайной величины Y, используемый в распределении Стьюдента, нормален, независим от совокупности $\{x_1,\ldots,x_n\}$ и также имеет интенсивность σ.

Далее угловыми скобками $<.>$ будем обозначать операцию нахождения математического ожидания относительно указываемых индексом случайных величин.

Теорема 1.

В принятых предположениях аддитивный функционал

$$\mathrm{J} = \frac{1}{n}\sum_{k=1}^{n} x_k^2 \tag{1}$$

имеет равновесную производящую функцию

$$Q_{\mathrm{J}}(\lambda; q, n) \equiv \left\langle \exp\left(-\lambda \mathrm{J}\right)\right\rangle_{\mathrm{J}} \tag{2}$$

следующего вида

$$Q_{\mathrm{J}}(\lambda; q, n) = \left(\frac{(1-q^2)R}{\left(a_+ - q^2\right)^2 a_+^{N-1} - \left(a_- - q^2\right)^2 a_-^{N-1}}\right)^{1/2}, \tag{3}$$

где

$$a_\pm = \frac{1}{2}\left(1 + q^2 + \frac{2\lambda\sigma^2}{n}(1-q^2) \pm R\right), \quad R = \left[\left((1+q^2) + \frac{\lambda\sigma^2}{n}\right)^2 - 4q^2\right]^{1/2} \tag{3a}$$

и $q = \exp(-\nu\tau)$ – коэффициент корреляции соседних отсчетов ОУ-процесса.

Плотность распределения вероятностей $f_{\mathrm{J}}(J)$ случайной величины J связана с производящей функцией $Q_{\mathrm{J}}(\lambda; q, n)$ обратным преобразованием Лапласа.

Таким образом, учет корреляции в указанной стратегии представляет практический интерес.

Важной особенностью теории оценивания и принятия решений является учет статистической связи между оценками уже известных распределений. В теории вероятностей и математической статистике известен аналог данному функционалу – это χ^2-распределение с числом степеней свободы равным $N-1$. Представляет интерес сравнить плотности распределения вероятностей рассматриваемого критерия качества (1) с χ^2-распределением. В работе [7] показано, что плотность распределения данного функционала совпадает с плотностью χ^2-распределения только в том случае, когда коэффициент корреляции $q = 0$, то есть корреляция между соседними сечениями случайного процесса отсутствует. В реальности, однако, значения случайных процессов в точках с соседними отсчетами являются в той или иной степени коррелированными.

Рассмотрим теперь случайную величину – стандартизованное отношение Стьюдента

$$\mathrm{T} = \mathrm{Y} / \sqrt{\mathrm{J}} = \frac{y_0}{\left(\frac{1}{n} \sum_{k=1}^{n} x_k^2 \right)^{1/2}} . \tag{5}$$

Статистический анализ случайной величины T проведем в терминах её характеристической функции $\Psi_{\mathrm{T}}(\xi)$, которую запишем в виде

$$\Psi_{\mathrm{T}}(\xi) = \left\langle \exp\left(i\xi \mathrm{Y} / \sqrt{\mathrm{J}} \right) \right\rangle_{\mathrm{Y,J}} , \tag{6}$$

здесь индексами указаны случайные величины, по которым необходимо произвести статистическое усреднение.

Из имеющихся двух случайных величин, пользуясь (4а), выполним усреднение по случайной величине Y:

$$\Psi_{\mathrm{T}}(\xi) = \int_{-\infty}^{\infty} \frac{dy_0}{\sqrt{2\pi}} \exp\left(-\frac{y_0^2}{2} \right) \left\langle \exp\left(i\xi\, y_0 / \sqrt{\mathrm{J}} \right) \right\rangle_{\mathrm{J}} = \left\langle \exp\left(-\frac{\xi^2}{2\mathrm{J}} \right) \right\rangle_{\mathrm{J}} .$$

Отсюда для плотности распределения $f_{\mathrm{T}}(t)$ случайной величины T

$$f_{\mathrm{T}}(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\xi \exp(i\xi t)\, \Psi_{\mathrm{T}}(\xi) , \tag{7}$$

имеем после интегрирования по ξ

$$f_{\mathrm{T}}(t) = \left\langle \sqrt{\frac{\mathrm{J}}{2\pi}} \exp\left(-t^2 \mathrm{J} / 2 \right) \right\rangle_{\mathrm{J}} ,$$

или, раскрывая оператор усреднения,

$$f_{\mathrm{T}}(t) = \int\limits_0^\infty \frac{dJ}{\sqrt{2\pi}} g_{\mathrm{J}}(J)\sqrt{J}\exp\left(-t^2 J/2\right),$$

где $g_{\mathrm{J}}(J)$ – плотность распределения вероятностей случайной величины J, отвечающая производящей функции $Q_{\mathrm{J}}(\lambda; q, n)$ (3),

$$g_{\mathrm{J}}(J) = \frac{1}{2\pi}\int_C d\lambda \exp(\lambda J)\, Q_{\mathrm{J}}(\lambda; q, n),$$

при этом контур интегрирования C выбирается таким образом, чтобы оставить все особенности функции $Q_{\mathrm{J}}(\lambda; q, n)$ слева. Подставляя это выражение в (7), находим после интегрирования по J

$$f_{\mathrm{T}}(t) = \frac{1}{2\pi i}\int_C d\lambda\, Q_{\mathrm{J}}(\lambda; q, n)(t^2 - 2\lambda)^{-3/2}. \tag{8}$$

Отсюда получим для интегральной функции распределения $F_{\mathrm{T}}(t)$ случайной величины T

$$F_{\mathrm{T}}(t) = \frac{1}{2\pi i}\int_C d\lambda\, Q_{\mathrm{J}}(\lambda; q, n)\int\limits_0^t d\tau(\tau^2 - 2\lambda)^{-3/2}.$$

Стандартное интегрирование приводит к выражению

$$F_{\mathrm{T}}(t) = \frac{1}{2\pi i}\int_C d\lambda\, Q_{\mathrm{J}}(\lambda; q, n)\, t/-2\lambda\sqrt{t^2 - 2\lambda}\ .$$

Используя теперь подстановки $\lambda = t^2 z/2$ и $z = 1 + \vartheta^2$, получим

$$F_{\mathrm{T}}(t) = -\frac{1}{4\pi i}\int_C d\lambda\, Q_{\mathrm{J}}(t^2 z/2; q, n)\frac{1}{z\sqrt{1-z}} = \frac{1}{2\pi i}\int_C d\vartheta\, Q_{\mathrm{J}}\left(\frac{1+\vartheta^2}{2}t^2; q, n\right)\frac{1}{1+\vartheta^2}.$$

При $t \ge 0$ контур интегрирования может быть выбран вдоль вещественной оси, что дает окончательно

$$F_{\mathrm{T}}(t) = \frac{1}{\pi}\int\limits_0^\infty d\vartheta\, Q_{\mathrm{J}}\left(\frac{1+\vartheta^2}{2}t^2; q, n\right)\frac{1}{1+\vartheta^2}. \tag{9}$$

Отсюда при $t \le 0$ найдем интегрированную функцию распределения с учетом четности полученного выражения.

Основные результаты. Для проверки правильности полученных формул необходимо предоставить плотности распределения при отсутствии корреляции, построенные с помощью формулы (8) и стандартным распределением Стьюдента. Данные плотности полностью совпадают, что свидетельствует о том, что полученные формулы найдены правильно.

Далее проанализируем влияние параметров на поведение плотности и функции распределения функционала (5). На рис. 1 представлены графики плотности и функции распределения функционала (5) при изменении параметра N.

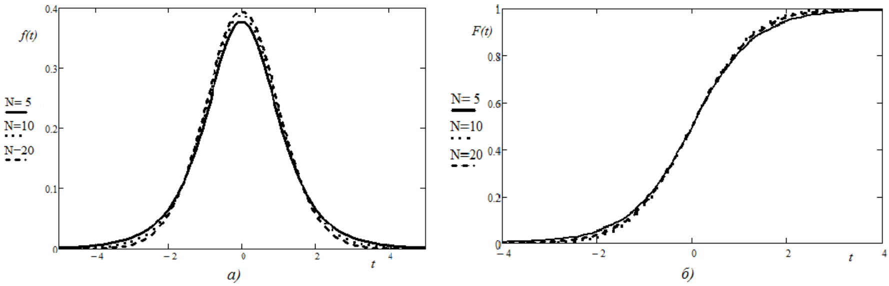

Рис. 1 – Графики плотности (*а*) и функции (*б*) распределения вероятностей для функционала (5) с параметрами: $N = 5, 10, 20$, $q = 0.5$, $\sigma = 1$

Анализируя рис. 1 можно сделать вывод о том, что они отвечают общим закономерностям, характерным для плотности и функции распределения Стьюдента при наличии корреляции.

На рис. 2-3 приведены плотности и функции распределения для функционала (5) соответственно.

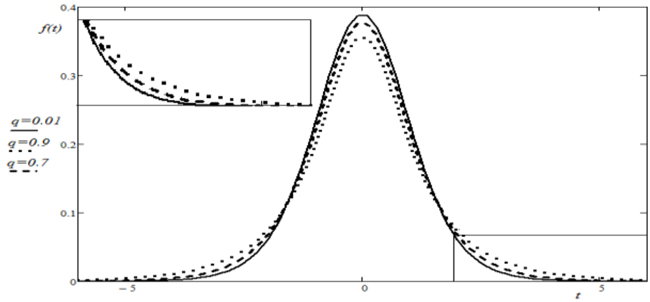

Рис. 2 – Графики плотности распределения вероятностей для функционала (5) с параметрами: $q = 0.01, 0.7, 0.9$, $N = 10$, $\sigma = 1$

С увеличением q график функции распределения имеет тенденцию локализоваться вокруг среднего значения функционала T. На рис. 2 – 3 эту закономерность можно пронаблюдать.

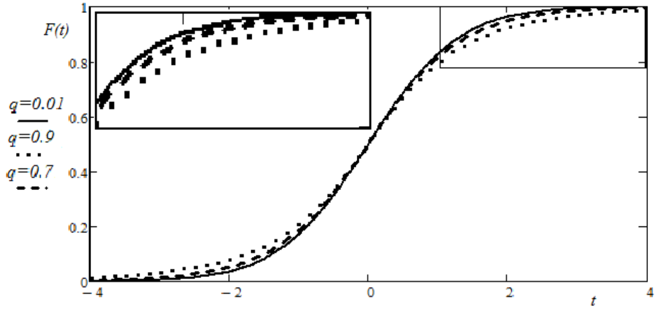

Рис. 3 – Графики функции распределения вероятностей для функционала (5) с параметрами: $q = 0.01, 0.7, 0.9$, $N = 10$, $\sigma = 1$

На основании полученных результатов (рис. 1-3) можно сделать вывод о том, что при увеличении корреляции функции плотности распределения закона Стьюдента становятся более пологими.

На практике представляет интерес поведение плотности и функции распределения функционала (5), именно, в периферийной области, т.к. на концах функции находятся пороговые значения, которые необходимы для контроля. Как можно наблюдать из рис. 2-3 в увеличенной области даже небольшие значения коэффициента корреляции q приводят к заметному изменению. Таким образом, учет корреляции в указанной стратегии представляет практический интерес.

Выводы. Полученные выше выражения справедливы для любых допустимых значений коррелятора q, в том числе и нулевых.

Найденные выражения для распределений стандартизованного отношения Стьюдента допускают предельный переход $n \to \infty$ для отсчетов случайного процесса на заданном интервале. Возможно также обобщение результатов на случаи комплекснозначных или многомерных случайных переменных.

Список литературы: **1.** *Гнеденко Б. В.* Курс теории вероятностей / *Б. В. Гнеденко.* – М. : Наука. – 1961. – 406 с. **2.** *Ивченко Г. И.* Математическая статистика / *Г. И. Ивченко, Ю. И. Медведев.* – М. : Высшая школа. – 1984. – 248 с. **3.** *Пугачев В. С.* Теория вероятностей и математическая статистика / *В. С. Пугачев.* – М. : Наука. – 1979. – 496 с. **4.** *Рытов С. М.* Введение в статистическую радиофизику / *С. М. Рытов.* – М. : Наука. – 1979. – 404 с. **5.** *Тихонов В. И.* Марковские процессы / *В. И. Тихонов, М. А. Миронов.* – М. : Сов. Радио. – 1977. – 488 с. **6.** *Мазманишвили А. С.* Континуальное интегрирование как метод решения физических задач / *А. С. Мазманишвили* – К. : Наукова думка. – 1987. – 224 с. **7.** *Сидоренко А. Ю.* Влияние параметров шероховатости на качество поверхностного слоя / *А. Ю. Сидоренко.* – Вісник НТУ «ХПІ». – 2004. – № 36. – С. 85–91.

Надійшла до редколегії 05.02.2012