УДК 519.2 Вестник СПбГУ. Сер. 1. Т. 2 (60). 2015. Вып. 3

О НЕКОТОРЫХ РЕГРЕССИОННЫХ СООТНОШЕНИЯХ, СВЯЗЫВАЮЩИХ ВЫБОРОЧНЫЕ СРЕДНИЕ И ПОРЯДКОВЫЕ СТАТИСТИКИ*

В. Б. Невзоров

Санкт-Петербургский государственный университет,
Российская Федерация, 199034, Санкт-Петербург, Университетская наб., 7/9

Получены характеризации некоторых вероятностных распределений регрессионными равенствами, в которых участвуют выборочные средние и разные элементы вариационного ряда. В качестве частного случая представлены соотношения, характеризующие t_2-распределение Стьюдента. Библиогр. 2 назв.

Ключевые слова: выборочные средние, порядковые статистики, характеризации вероятностных распределений, семейство t_2-распределений Стьюдента.

Введение. Рассмотрим последовательность независимых случайных величин (с. в.) $X, X_1, X_2, \ldots$ с общей функцией распределения (ф. р.) $F(x)$ и порядковые статистики $X_{1,n} \le X_{2,n} \le \ldots \le X_{n,n}$, соответствующие с. в. $X_1, X_2, \ldots, X_n, n = 1, 2, \ldots$ Пусть $X(n) = \frac{X_1+X_2+\ldots+X_n}{n}$, $n = 1, 2, \ldots$, обозначает выборочные средние.

В связи с исследованием свойств распределений Стьюдента в работе [2] был получен следующий результат.

Теорема 1. *Пусть $X_1, X_2, \ldots$ — независимые одинаково распределенные случайные величины с конечным математическим ожиданием, функцией распределения $F(x)$ и плотностью распределения $f(x)$. Пусть также $f(x) > 0$, если $\gamma < x < \delta$, где $\gamma = \inf\{x : F(x) > 0\}$ и $\delta = \sup\{x : F(x) < 1\}$. Тогда эквивалентными являются следующие утверждения:*

$$a)\ E(X(3)|X_{2,3} = x) = x, \quad \gamma < x < \delta; \tag{1}$$

$$b)\ E\left(\frac{X_{1,3} + X_{3,3}}{2}|X_{2,3} = x\right) = x, \quad \gamma < x < \delta; \tag{2}$$

$$c)\ (F(x)(1 - F(x))^{\frac{3}{2}} = Cf(x), \quad \gamma < x < \delta, \tag{3}$$

где $C > 0$ — некоторая константа;

d) $\gamma = -\infty, \quad \delta = \infty$ и $F(x) = F_2(\frac{x-\mu}{\sigma}), \quad -\infty < \mu < \infty, \quad \sigma > 0$, где

$$F_2(x) = \frac{1 + \frac{x}{(2+x^2)^{\frac{1}{2}}}}{2}, \quad -\infty < x < \infty. \tag{4}$$

По сути дела этот результат можно рассматривать как две разные характеризации семейства распределений Стьюдента с двумя степенями свободы, представители которого имеют ф. р. $F_2(\frac{x-\mu}{\sigma})$. Одна характеризация основана на регрессионном соотношении (1), связывающем выборочное среднее $X(3)$ и порядковую статистику (выборочную медиану) $X_{2,3}$. В другом регрессионном соотношении (2) фигурируют середина размаха $\frac{X_{1,3}+X_{3,3}}{2}$ и та же порядковая статистика $X_{2,3}$.

*Работа выполнена при финансовой поддержке СПбГУ (грант № 6.38.672.2013) и РГНФ (грант № 13-02-00338).

В этой же работе [2] исследовано более общее, чем (2), соотношение. Для нечетных $n = 2k+1, k = 1, 2, \ldots$, были введены середины квазиразмахов

$$M(k) = \frac{X_{k,2k+1} + X_{k+2,2k+1}}{2}, \quad k = 1, 2, \ldots .$$

В частности, при $k = 1$ с. в. $M(1)$ совпадает с серединой размаха $\frac{X_{1,3}+X_{3,3}}{2}$, рассмотренной в теореме 1.

Пусть $G(x)$ — функция, обратная $F(x)$. Было показано, что справедливы следующие результаты.

Теорема 2. *В условиях теоремы 1 для любого* $k = 1, 2, \ldots$ *эквивалентны следующие утверждения:*

$$a)\ E(M(k)|X_{k+1,2k+1} = x) = x, \quad \gamma < x < \delta; \tag{5}$$

$$b)\ (F(x)(1 - F(x)))^{\frac{k}{2}+1} = Cf(x), \quad \gamma < x < \delta; \tag{6}$$

где $C = C(k)$ *— некоторая положительная константа;*

c) $\gamma = -\infty, \quad \delta = \infty$ *и*

$$G(x) = C \int_{\frac{1}{2}}^{x} (u(1-u))^{-(\frac{k}{2}+1)} du + d, \quad 0 < x < 1, \tag{7}$$

где $C > 0$, $-\infty < d < \infty$.

Очевидно, что при $k = 1$ утверждения теоремы 2 совпадают с соответствующими утверждениями теоремы 1.

Основные результаты. Интересно выяснить, какие обобщения можно получить, рассматривая уже при любом n регрессионные равенства вида (1), связывающие выборочные средние $X(n)$ и произвольные порядковые статистики $X_{k,n}$.

Справедлив следующий результат.

Теорема 3. *В условиях теоремы 1 для любых* $n = 3, 4, \ldots$ *и* $1 < k < n$ *эквивалентны следующие утверждения:*

$$a)\ E(X(n)|X_{k,n} = x) = x; \tag{8}$$

$$b)\ (F(x))^{2-\beta}(1 - F(x))^{1+\beta} = Cf(x), \quad \gamma < x < \delta, \tag{9}$$

где $\beta = \frac{k-1}{n-1}$, *а* $C = C(k, n)$ *— некоторая положительная константа;*

c) $\gamma = -\infty, \quad \delta = \infty$ *и*

$$G(x) = C \int_{\frac{1}{2}}^{x} u^{-(1+\beta)}(1-u)^{-(2-\beta)} du + d, \quad 0 < x < 1, \tag{10}$$

где $C = C(\beta) > 0$, $-\infty < d < \infty$ *— некоторые константы.*

Замечание 1. Условие $1 < k < n$ в теореме 3 существенно, так как очевидно, что $P\{X_{1,n} < X(n) < X_{n,n}\} = 1$.

Замечание 2. Видим, сравнивая соотношения (6) и (9), что распределения, характеризуемые равенствами (5) и (8), могут совпадать лишь в случае, когда $\beta = \frac{1}{2}$ в теореме 3, а $\frac{k}{2} + 1 = \frac{3}{2}$ в теореме 2. Это соответствует значению $k = 1$ в равенстве (6) и значениям $n = 2k - 1$, $k = 2, 3, \ldots$, в теореме 3. Получаем, что единственный вариант, когда совпадают распределения, характеризуемые соотношениями (5) и (8), представлен в теореме 1.

Замечание 3. Если $\beta = \frac{1}{2}$ в (9), что, как было отмечено выше, выполняется, если $n = 2k-1, k = 2, 3, \ldots$, то регрессионное соотношение (8) характеризует уже знакомое нам по теореме 1 семейство t_2-распределений Стьюдента.

Доказательство теоремы 3. Справедливо очевидное равенство

$$E(X(n)|X_{k,n} = x) = E(X_1|X_{k,n} = x). \tag{11}$$

Естественно, что вместо X_1 в правую часть (11) можно подставить любую из с. в. $X_2, X_3, \ldots, X_n$. Известно (см., например, [1] или [2]), что

$$E(X_1|X_{k,n} = x) = \frac{x}{n} + \frac{k-1}{n}E(X|X < x) + \frac{n-k}{n}E(X|X > x). \tag{12}$$

В условиях теоремы 3 приходим к равенству

$$x = E(X(n)|X_{k,n} = x) = E(X_1|X_{k,n} = x) = \\ = \frac{x}{n} + \frac{k-1}{nF(x)}\int_{-\infty}^{x} udF(u) + \frac{n-k}{n(1-F(x)}\int_{x}^{\infty} udF(u). \tag{13}$$

Учитывая, что

$$\int_{-\infty}^{x} udF(u) = xF(x) - \int_{-\infty}^{x} F(u)du$$

и

$$\int_{x}^{\infty} udF(u) = x(1-F(x)) + \int_{x}^{\infty} (1-F(u))du,$$

получаем, что из (13) следует соотношение

$$\frac{n-k}{1-F(x)}\int_{x}^{\infty} (1-F(u))du = \frac{k-1}{F(x)}\int_{-\infty}^{x} F(u)du, \quad \gamma < x < \delta. \tag{14}$$

Равенство (14) можно переписать в виде

$$\frac{1-F(x)}{\int_{x}^{\infty} (1-F(u))du} = \frac{\alpha F(x)}{\int_{-\infty}^{x} F(u)du},$$

где $\alpha = \frac{n-k}{k-1}$, что равносильно соотношению

$$\left(-\ln\int_{x}^{\infty} 1-F(u)du\right)' = \left(\alpha\ln\int_{-\infty}^{x} F(u)du\right)'. \tag{15}$$

Из (15) получаем, что

$$\int_{x}^{\infty} (1-F(u))du = \frac{C_1}{(\int_{-\infty}^{x} F(u)du)^{\alpha}}, \tag{16}$$

где C_1 — некоторая константа.

После дифференцирования в равенстве (16) приходим к соотношению

$$(1 - F(x)) = \alpha C_1 F(x) \left(\int_{-\infty}^{x} F(u) du \right)^{-(\alpha+1)},$$

которое можно переписать в виде

$$\int_{-\infty}^{x} F(u) du = C_2 \left(\frac{F(x)}{1 - F(x)} \right)^{1/(\alpha+1)},$$

где $C_2 = (\alpha C_1)^{1/(\alpha+1)} > 0$.

Еще одно дифференцирование приводит к равенству

$$F(x) = \frac{C_2}{\alpha+1} f(x) (F(x))^{1/(\alpha+1)-1} (1 - F(x))^{-(1/(\alpha+1)+1)},$$

из которого получаем, введя обозначение

$$\beta = \frac{1}{\alpha+1} = \frac{k-1}{n-1},$$

что

$$(F(x))^{2-\beta} (1 - F(x)^{1+\beta}) 1 + \beta = C f(x), \quad \gamma < x < \delta,$$

где C (при фиксированных значениях k и n) — некоторая положительная константа. Отметим также, что $0 < \beta < 1$, поскольку $1 < k < n$. Соотношение (9) доказано.

Если учесть, что $G'(x) = \frac{1}{f(G(x))}$, где $G(x)$ — функция, обратная $F(x)$, то, подставляя $G(x)$ вместо x в соотношение (9), приходим к равенству

$$\frac{C}{G'(x)} = x^{2-\beta} (1 - x)^{1+\beta}, \quad 0 < x < 1,$$

из которого получаем , что $G(x)$ можно представить в виде

$$G(x) = C \int_{\frac{1}{2}}^{x} u^{-(1+\beta)} (1 - u)^{-(2-\beta)} du + d, \quad 0 < x < 1,$$

где d — произвольная константа, т. е. справедливо утверждение (10).

Остается добавить, что поскольку $0 < x < 1$, из соотношения (10) следует, что $\gamma = G(0) = -\infty$ и $\delta = G(1) = \infty$. Теорема доказана.

Литература

1. *Ahsanullah M., Nevzorov V. B.* On characterizations based on record values and order statistics // J. Statistical Planning and Inference, 1997. Vol. 63. P. 271–289.

2. *Невзоров В. Б.* Об одном свойстве распределения Стьюдента с двумя степенями свободы // Записки научных семинаров ПОМИ, 2002. Т. 294. С. 148–157.

Статья поступила в редакцию 26 марта 2015 г.

Сведения об авторе

Невзоров Валерий Борисович — доктор физико-математических наук, профессор; probabil@pisem.net

ON SOME REGRESSIONAL RELATIONS CONNECTING SAMPLE MEANS AND ORDER STATISTICS

Valery B. Nevzorov

St.Petersburg State University, Universitetskaya nab., 7/9, St.Petersburg, 199034, Russian Federation; probabil@pisem.net

New characterizations of some probability distributions based on regressional properties of sample means and order statistics are obtained. Some relations, which characterize Student's t_2-distribution, are presented as a partial case. Refs 2.

Keywords: sample means, order statistics, characterizations of probability distributions, Student's t_2-distribution.

References

1. Ahsanullah M., Nevzorov V. B., "On characterizations based on record values and order statistics", *J. Statistical Planning and Inference* **63**, 271–289 (1997).

2. Nevzorov V. B., "On one property of Student's distribution with two degrees of freedom", *Zapiski Nauchnykh Seminarov POMI* **294**, 148–157 (2002) (in Russian). (English translation: *J. Math. Sci.* **127**(1), 1757–1762 (2005).)