

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ
РОССИЙСКОЙ ФЕДЕРАЦИИ
ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

МАТЕРИАЛЫ

**Первой Всероссийской молодежной
научной конференции
«МАТЕМАТИЧЕСКОЕ
И ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ
ИНФОРМАЦИОННЫХ,
ТЕХНИЧЕСКИХ
И ЭКОНОМИЧЕСКИХ СИСТЕМ»**

Томск, 17–18 мая 2013 г.

*Под общей редакцией
кандидата технических наук И.С. Шмырина*

Томск
Издательский Дом Томского государственного университета
2013

Для рядов температуры на алтайских станциях такие же результаты получены и при исключении сезонной составляющей, из чего можно сделать вывод о сходной динамике сезонности.

ЛИТЕРАТУРА

1. Корн Г. Справочник по математике для научных работников и инженеров: определения, теоремы, формулы. / Г. Корн, Т. Корн. – М. : Главная редакция физико-математической литературы издательства «Наука», 1974. – 832 с.
2. Суетин П. К. Классические ортогональные полиномы / П. К. Суетин. – 2-е изд., доп. – М. : Главная редакция физико-математической литературы издательства «Наука», 1979. – 416 с.
3. Бухановский А. В. Вероятностный анализ и моделирование вертикально неоднородного океанологического поля / А.В. Бухановский, Е.Н. Захарченко, В.А. Рожков // Навигация и гидрография. – 1999. - № 9. – С.86-103.
4. Данилов Ю. А. Многочлены Чебышева / Ю. А. Данилов. – Минск : Вышэйшая школа, 1984. – 157 с.
5. Колемаев В. А. Теория вероятностей и математическая статистика / В.А. Колемаев, В.Н. Калинина. – М.: ИНФРА-М, 1997. – 302 с.

ИССЛЕДОВАНИЕ ЦЕНЫ ПОТРЕБИТЕЛЬСКОГО СПРОСА НА ТОВАР-НОВИНКУ ПО ДВУКРАТНО СПРАВА ЦЕНЗУРИРОВАННЫМ ДАННЫМ

М. А. Муравлева

Томский государственный университет
E-mail: marmyr90@rambler.ru

В процессе проведения маркетинговых исследований маркетологи постоянно сталкиваются с неполными наблюдениями, когда значение исследуемого показателя известно с точностью до некоторого интервала или даже пропущено. При этом зачастую обработка подобных данных проводится без учета их природы, к ним нередко применяются классические методы, ориентированные на полное знание каждого наблюдения. Полученные в итоге результаты являются некорректными и не позволяют сделать адекватные выводы. К неполным, цензурированным данным необходимо применять специально разработанные для этих целей модели [1], которые существенно повышают точность обработки [2]. Еще большее значение при этом придается знанию любой априорной информации об исследуемой случайной величине (с.в.), так как ее привлечение способствует повышению качества статистических процедур [2].

В данной работе рассматривалась случайная величина $\tau > 0$ — цена в рублях за единицу товара — с функцией распределения (ф.р.) $F(t) = P(\tau \leq t)$. Прогрессивно цензурированная выборка (ц.в.) объема N $(X, I) = \{(X_1, I_1), \dots, (X_N, I_N)\}$, где для $i = \overline{1, N}$

$$I_i = \begin{cases} 0, & X_i - \text{полное наблюдение;} \\ 1, & X_i - \text{наблюдение до цензурирования,} \end{cases}$$

строилась по следующей схеме: цензурирование интервалом I типа, моменты цензурирования T_1, T_2 — не случайны, количество неполных наработок в интервале $[T_1, T_2]$ — с.в., численно равная доле g , $0 < g < 1$, от суммарного числа наблю-

даемых объектов в конце интервала $[0, T_1]$ и в начале интервала $[T_2, T]$, т. е. доле от $(N - r_1 - r_2)$ — количества оставшихся под наблюдением объектов в интервале $[T_1, T_2]$.

Тогда оценка функции распределения определяется формулой [1]

$$F_N^u(t) = \begin{cases} 0, & t \leq 0 \\ \frac{1}{N} \sum_{i=1}^N I_{[0,t)}(X_i) \bar{I}_i, & 0 \leq t \leq T_1 \\ \left. \begin{aligned} & \frac{r_1}{N} + \frac{1}{(1-g)N} \sum_{i=1}^N I_{[T_1,t)}(X_i) \bar{I}_i, N_1 > 0 \\ & \frac{r_1}{N} \end{aligned} \right\} & T_1 \leq t \leq T_2 \quad (1) \\ 1 - \frac{r_2}{N} + \frac{1}{N} \sum_{i=1}^N I_{[T_2,t)}(X_i) \bar{I}_i, & T_2 \leq t \leq T \\ 1, & t > T \end{cases}$$

где для $i = \overline{1, N}$ $\bar{I}_i = 1 - I_i$, r_1 — число полных наработок в интервале $[0, T_1]$, r_2 — число полных наработок в интервале $[T_2, T]$, $N_1 = (N - r_1 - r_2)(1 - g)$.

Оценка (1) является непараметрической и асимптотически несмещенной [2], в [1] было показано, что при $N_1 > 0$ подобные модели соответствуют оценке Каплана-Мейера.

Найдем оценку среднего значения цены методом подстановки [3]. Получим

$$\bar{x}_N^u = \int_0^{\infty} x dF_N^u(x) = \frac{1}{N} \sum_{i=1}^{r_1} X_i + \frac{1}{(1-g)N} \sum_{i=r_1+1}^{N-r_2-1} X_i + \frac{1}{N} \sum_{i=r_2}^N X_i. \quad (2)$$

Оценка (2) также является асимптотически несмещенной.

Приведенные формулы использовались при анализе желаемой цены на продукт для прослушивания звуковых файлов, за которую потребители готовы приобрести данный товар. Опрашивались студенты Томского государственного университета (ТГУ), объем выборки $N = 61$, моменты цензурирования $T_1 = 100$ и $T_2 = 500$. Для большего удобства все данные были масштабированы в пропорции 1:100.

По формуле (1) строилась оценка $F_N^u(x)$, статистические свойства которой изучались с помощью бутстреп-метода с параметром моделирования $M = 10000$. В частности, анализировалось математическое ожидание $F_N^{u\delta}(x) = MF_N^u(x)$ (рис. 1).

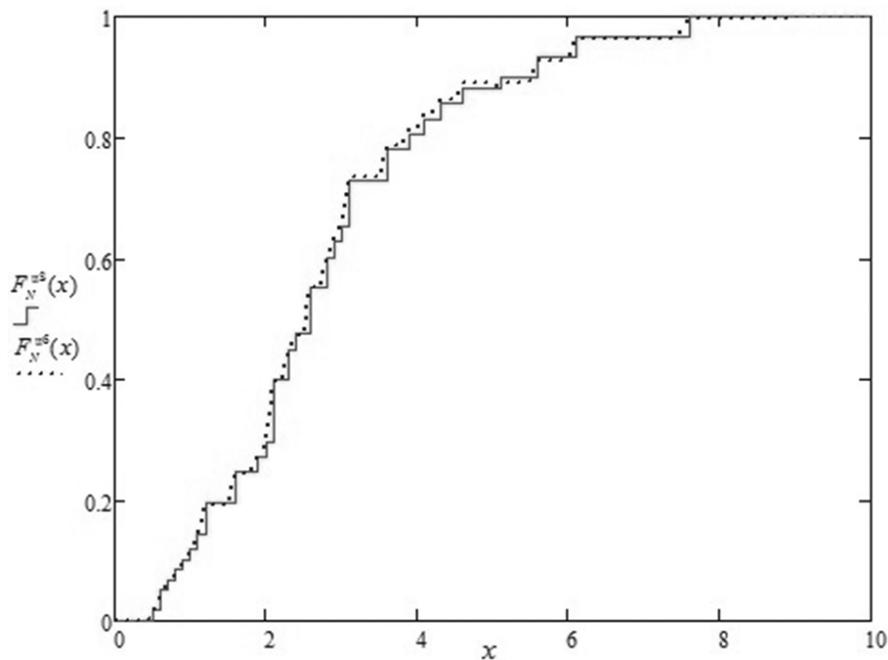


Рис. 1. Графики оценки ф. р. $F_N^{\text{II}}(x)$ и $F_N^{\text{IIб}}(x)$ для данной выборки
 $N = 61, T_1 = 1, T_2 = 5, M = 10000$

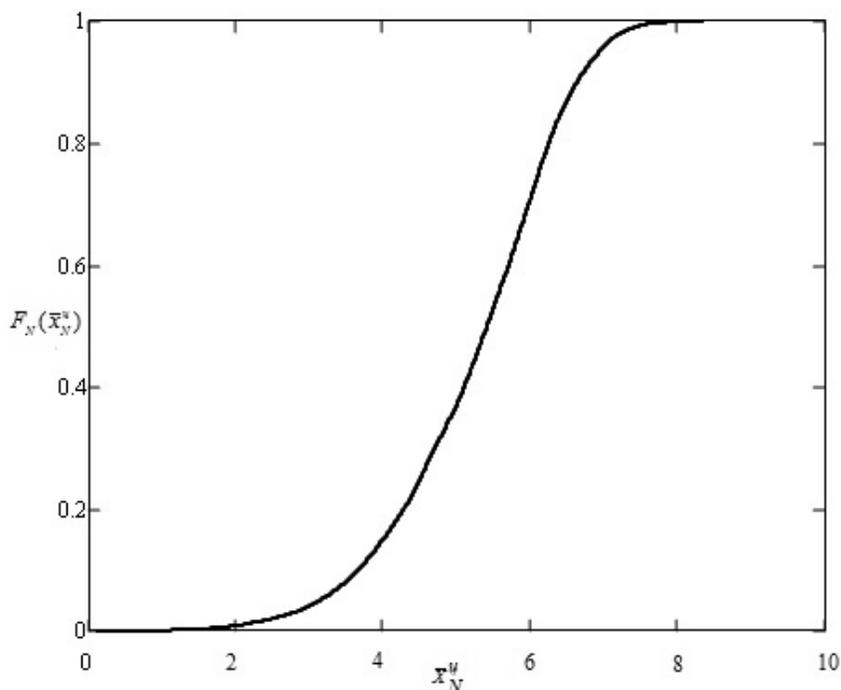


Рис. 2 График оценки ф.р. $F_N(\bar{x}_N^{\text{II}})$ для данной выборки
 $N = 61, T_1 = 1, T_2 = 5, M = 10000$

Если бы маркетологи действовали по классической схеме, то для построения доверительного интервала традиционно использовался бы нормальный закон, что привело бы к следующему результату: с вероятностью 95% средняя цена, которую готовы заплатить студенты ТГУ, попадет в интервал $[2.952; 2.984]$ или в первоначальных обозначениях $[295.2; 298.4]$ руб./шт.

Однако использование методов обработки цензурированных данных привело к несколько иным выводам. С помощью формулы (2) было получено, что средняя цена, за которую готовы приобрести товар студенты ТГУ, $\bar{X}_N^n = 296.9$ руб./шт.

С помощью бутстреп-метода моделировалось распределение \bar{X}_N^n при параметре моделирования $M = 10000$. Результат работы алгоритма приведен на рис. 2. Бутстреп-распределение позволило получить доверительный интервал для средней желаемой цены. Так, для студентов ТГУ, с вероятностью 95% средняя запрашиваемая цена попадает в интервал $[2.963; 2.974]$ или в первоначальных обозначениях $[296.3; 297.4]$ руб./шт.

Таким образом, использование бутстреп-метода, не требующего информации о виде распределения исследуемых данных, позволило улучшить статистические свойства оценок неизвестной ф.р. при работе с цензурированными данными. Средняя запрашиваемая цена с учетом пожеланий студентов Томского государственного университета оказалась в более узких пределах, чем при традиционных методах исследования. Такой результат позволит продавцу вывести новый товар на рынок с уточненной ценой.

ЛИТЕРАТУРА

1. Анализ надежности технических систем по цензурированным выборкам / В. М. Скрипник, А. Е. Назин, Ю. Г. Приходько, Ю. Н. Благовещенский. — М.: Радио и связь, 1988. — 184 с.: ил.
2. Зенкова Ж. Н. Статистическая обработка данных с учетом симметрии распределения. — Германия: LAP LAMBERT Academic Publishing GmbH&Co., 2011. — 181 с.
3. Кобзарь А. И. Прикладная математическая статистика. Для инженеров и научных работников. / А. И. Кобзарь — М.: Физматлит, 2006. — 816 с.

СИНТЕЗ ИНДИКАТОРНЫХ ПРОЦЕДУР ДЛЯ МОДЕЛИ МНОГОКВАНТИЛЬНОЙ РЕГРЕССИИ С ГЕТЕРОСКЕДАСТИЧНОСТЬЮ И МАРКОВСКИМИ ОСТАТКАМИ

Л. Ю. Пасечникова

Томский государственный университет
E-mail: pasechnikova_lyu@mail.ru

Целью данной работы является синтез процедур знакового статистического анализа модели с квантильной регрессией и Марковскими ошибками. В работе использованы результаты, ранее полученные для случая знакового анализа независимых наблюдений в [1], индикаторного анализа независимых наблюдений в [6], наблюдений, сопровождающихся Марковскими остатками в [5], и наблюдений, сопровождающихся явлением гетероскедастичности в [7].