

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

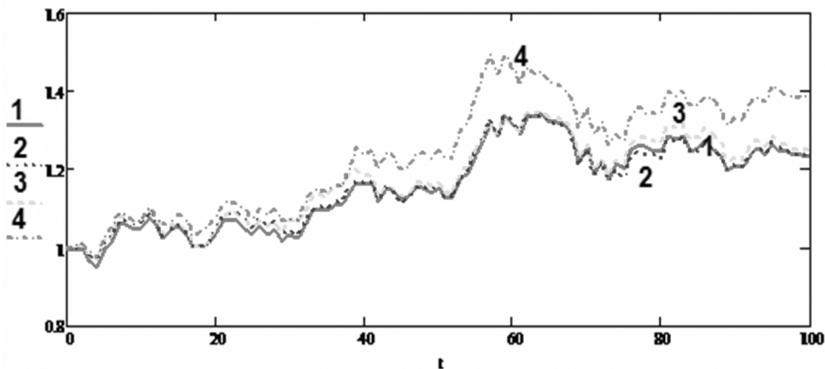


Рис. 4. Влияние изменений уровней значимости на динамику капитала инвестиционного портфеля, уровни значимости: 1 — $\beta = 0.25$ и $\alpha = 0.75$, 2 — $\beta = 0.10$ и $\alpha = 0.90$, 3 — $\beta = 0.05$ и $\alpha = 0.95$, 4 — $\beta = 0.01$ и $\alpha = 0.99$

В работе проведён экспериментальный анализ инвестиционных портфелей, построенных по двум критериям, а также анализ влияния длины исторического горизонта и уровней значимости в критерии отношения квантилей (VR) на динамику инвестиционного портфеля.

ЛИТЕРАТУРА

1. Sharpe W. F. The Sharpe Ratio // J. Portfolio Management. 1994.
2. Шарп У. Ф. Инвестиции / У. Ф. Шарп, Г. Дж. Александер, Дж. В. Бэйли. — М.: Инфра-М, 2001. — 1028 с.
3. Rachev S. T. Advanced stochastic models, risk assessment, and portfolio optimization: The ideal risk, uncertainty, and performance measures / S. T. Rachev, S. V. Stoyanov, F. J. Fabozzi. — USA.: Wiley. — 361 p.
4. Кибзун А. И. Сравнение критериев Var и CVaR / А. И. Кибзун, Е. А. Кузнецов // Автоматика и телемеханика. — 2003. — №7. — С. 153–163.

АНАЛИЗ ЕЖЕМЕСЯЧНОГО ДОХОДА ПО ЦЕНЗУРИРОВАННЫМ СПРАВА ДАННЫМ

Е. В. Рахимова

Томский государственный университет
E-mail: rakhimova2010@mail.ru

Цензурированные данные часто встречаются при проведении маркетинговых исследований, в частности, при опросах и анкетировании, когда опрашиваемый не может или не хочет указывать конкретное значение исследуемого показателя, например, своего дохода. Нередко маркетологи обрабатывают полученные цензурированные выборки некорректно, игнорируя цензурированные значения или ориентируясь на границы интервалов, что в итоге приводит к неадекватным результатам, которые используются как основа при принятии управленческих решений, чреватых финансовыми потерями.

В данной работе рассматривалась с.в. $\tau > 0$ — ежемесячный доход в рублях — с 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 справа I типа, моменты цензурирования $T_j, j = \overline{1, m}, m = 2$ — не случайны, r_j — число полных наработок в интервале $[T_{j-1}, T_j]$, n_j — число неполных наработок в интервале $[T_{j-1}, T_j]$, $\sum_{j=1}^m (r_j + n_j) = N, N_i = N - \sum_{j=0}^{i-1} (r_j + n_j)$ — число оставшихся объектов, за которыми ведется наблюдение в промежутке $(T_{i-1}, T_i]$.

В случае $m = 2$ оценка ф.р. $F(t)$ будет иметь вид [1]:

$$F_N^u(t) = \begin{cases} 0, & t \leq 0, \\ \frac{1}{N} \sum_{i=1}^N I_{[0,t)}(X_i) \bar{I}_i, & t \in (0, T_1], \\ \left. \begin{aligned} & \frac{r_1}{N} + \frac{1 - \frac{r_1}{N}}{N - r_1 - n_1} \sum_{i=1}^N I_{[T_1,t)}(X_i) \bar{I}_i, & N_1 > 0, \\ & \frac{r_1}{N}, & N_1 \leq 0, \end{aligned} \right\} & t \in (T_1, T_2], \\ \left. \begin{aligned} & \frac{r_1}{N} + \frac{r_2 \left(1 - \frac{r_1}{N}\right)}{N - r_1 - n_1} + \frac{1 - \frac{r_1}{N} - \frac{r_2 \left(1 - \frac{r_1}{N}\right)}{N - r_1 - n_1}}{N - r_1 - n_1 - r_2 - n_2} \sum_{i=1}^N I_{[T_1,t)}(X_i) \bar{I}_i, & N_2 > 0, \\ & \frac{r_1}{N} + \frac{r_2 \left(1 - \frac{r_1}{N}\right)}{N - r_1 - n_1}, & N_2 \leq 0, \end{aligned} \right\} & t \in (T_2, T], \\ 1, & t \leq 0. \end{cases} \quad (1)$$

Индекс "ц" здесь — сокращение от цензурированный. Оценка (1) является асимптотически несмещенной.

Найдем оценку среднего значения дохода методом подстановки [2]. Получим

$$\begin{aligned} \bar{X}_N^u = \int_0^{+\infty} x dF_N^u(x) &= \frac{1}{N} \sum_{i=1}^N X_i \cdot \bar{I}_i \cdot I_{[0,T_1)}(X_i) + \frac{1 - \frac{r_1}{N}}{N - r_1 - n_1} \sum_{i=1}^N X_i \cdot \bar{I}_i \cdot I_{[T_1,T_2)}(X_i) + \\ &+ \frac{1 - \frac{r_1}{N} - \frac{r_2 \left(1 - \frac{r_1}{N}\right)}{N - r_1 - n_1}}{N - r_1 - n_1 - r_2 - n_2} \sum_{i=1}^N X_i \cdot \bar{I}_i \cdot I_{[T_2,T]}(X_i), \end{aligned} \quad (2)$$

Данная оценка также является асимптотически несмещенной.

Приведенные формулы использовались при анализе ежемесячного дохода студентов двух вузов г. Томска: Томского государственного университета (ТГУ) и Сибирского государственного медицинского университета (СибГМУ), при этом в обоих вузах объем выборок $N = 80$, моменты цензурирования $T_1 = 2000$ и $T_2 = 5000$. Опрос проводился в ноябре 2011 г.

На рис. 1 и 2 показаны графики функций $F_N^u(t)$.

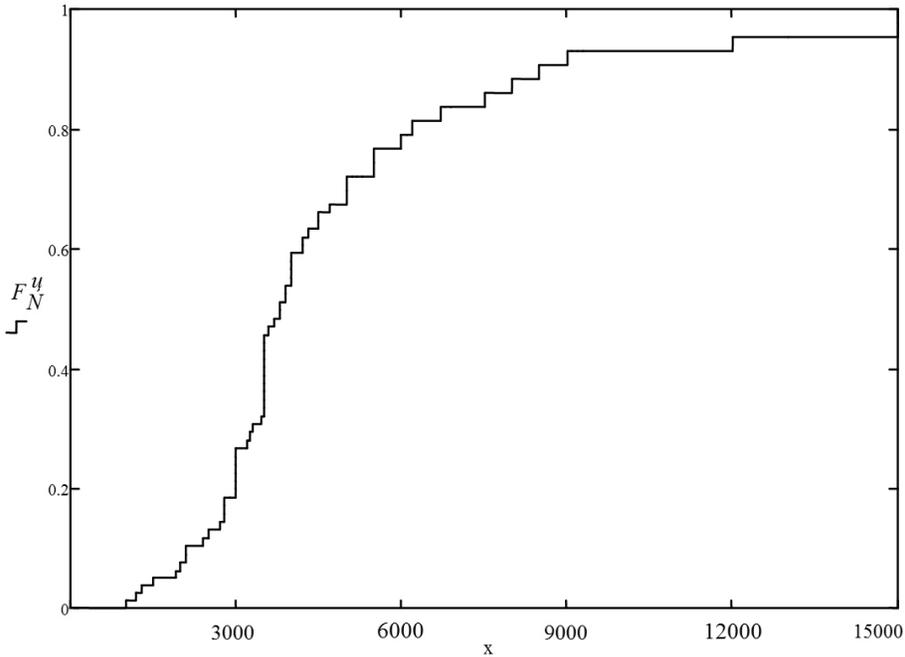


Рис. 1. График оценки ф.р. $F_N^u(t)$ для данных о доходах студентов ТГУ;

$$N = 80, T_1 = 2000, T_2 = 5000, n_2 = 6, n_3 = 10$$

С помощью формулы (2) было получено, что средний ежемесячный доход студентов, учащихся в ТГУ, $\bar{X}_N^u \approx 4835.59$ руб./мес., в СИБГМУ $\bar{X}_N^u \approx 4123.95$ руб./мес.

Статистические свойства полученных оценок исследовались с помощью бутстреп-метода [3]. В частности, анализировалось математическое ожидание $F_N^{ub}(t) = MF_N^u(t)$ для каждой выборки. Параметр моделирования $M = 10000$. Результаты бутстреп-моделирования приведены на рис. 3 и 4.

Также с помощью бутстреп-метода моделировалось распределение \bar{X}_N^u . Результат работы алгоритма приведен на рис. 5 и 6 при параметре моделирования $M = 10000$.

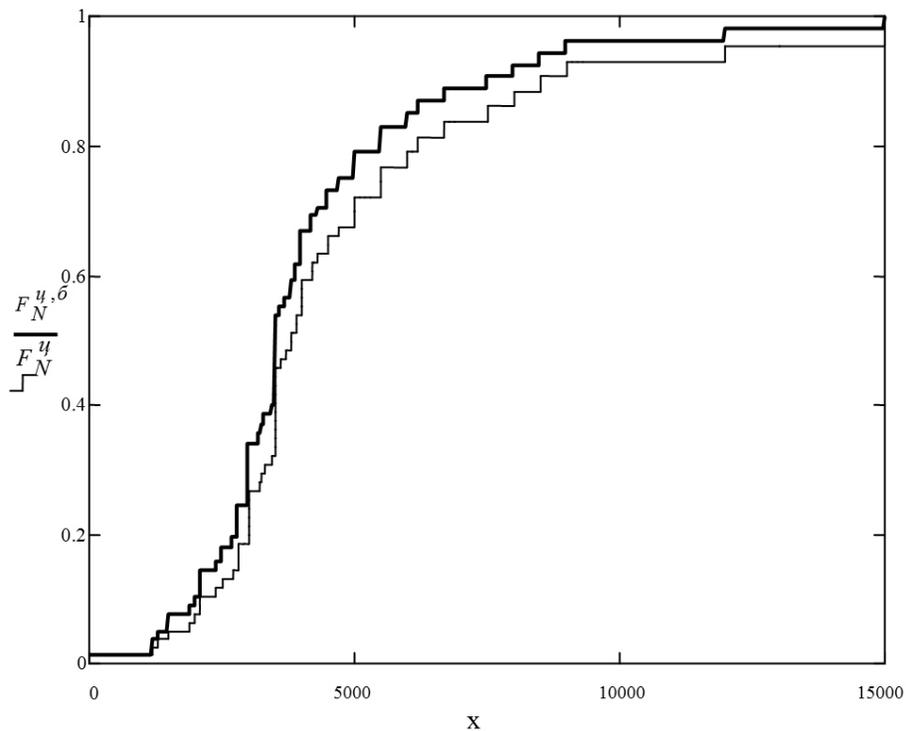


Рис. 2. График оценки ф.р. $F_N^u(t)$ для данных о доходах студентов СибГМУ;

$$N = 80, T_1 = 2000, T_2 = 5000, n_2 = 11, n_3 = 5$$

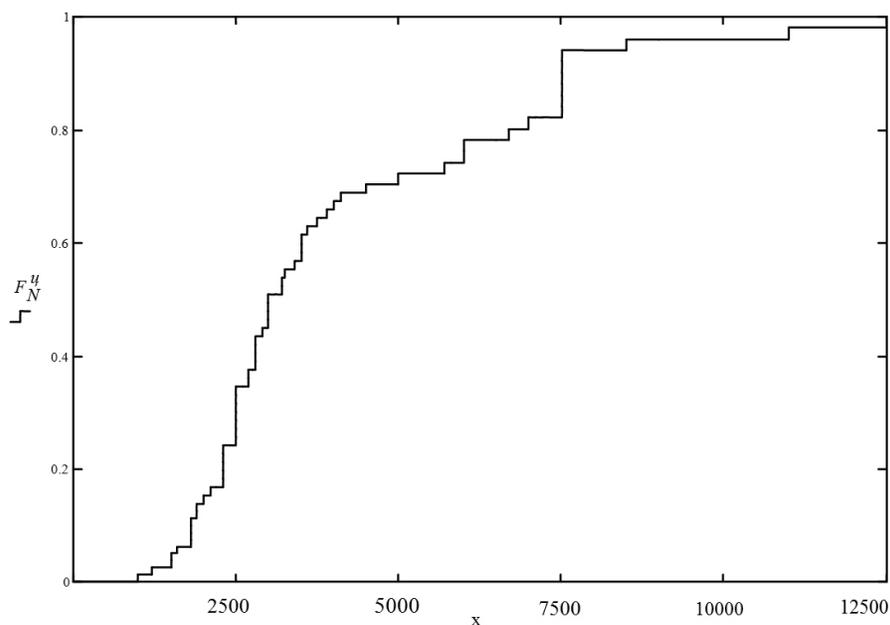


Рис. 3. Графики оценки ф.р. $F_N^u(t)$ и $F_N^{u\hat{\sigma}}(t)$ для данных о доходах студентов ТГУ;

$$N = 80, T_1 = 2000, T_2 = 5000, n_2 = 6, n_3 = 10$$

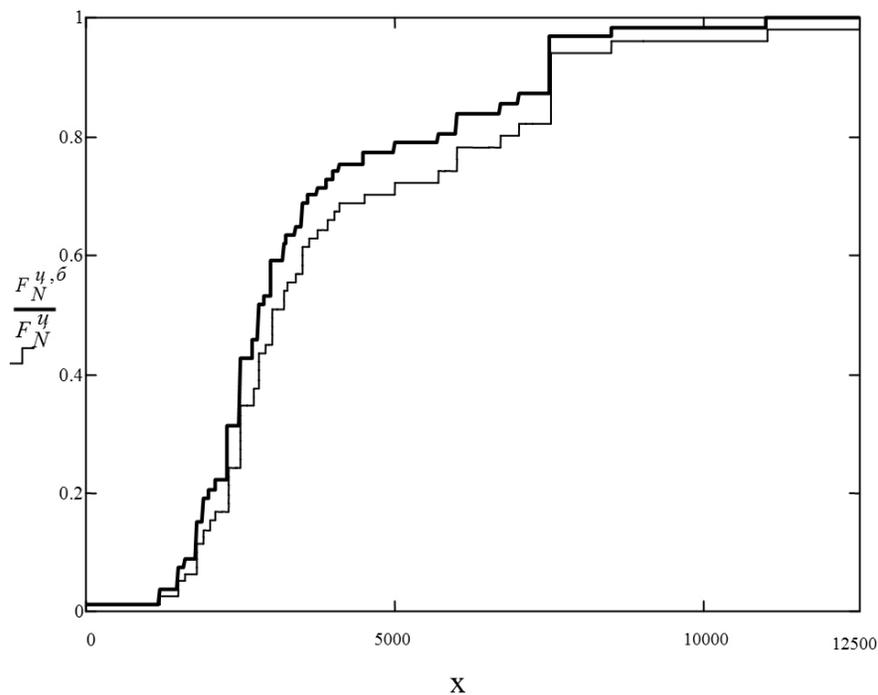


Рис. 4. Графики оценки ф.р. $F_N^u(t)$ и $F_N^{u\hat{\sigma}}(t)$ для данных о доходах студентов СибГМУ;
 $N = 80$, $T_1 = 2000$, $T_2 = 5000$, $n_2 = 11$, $n_3 = 5$

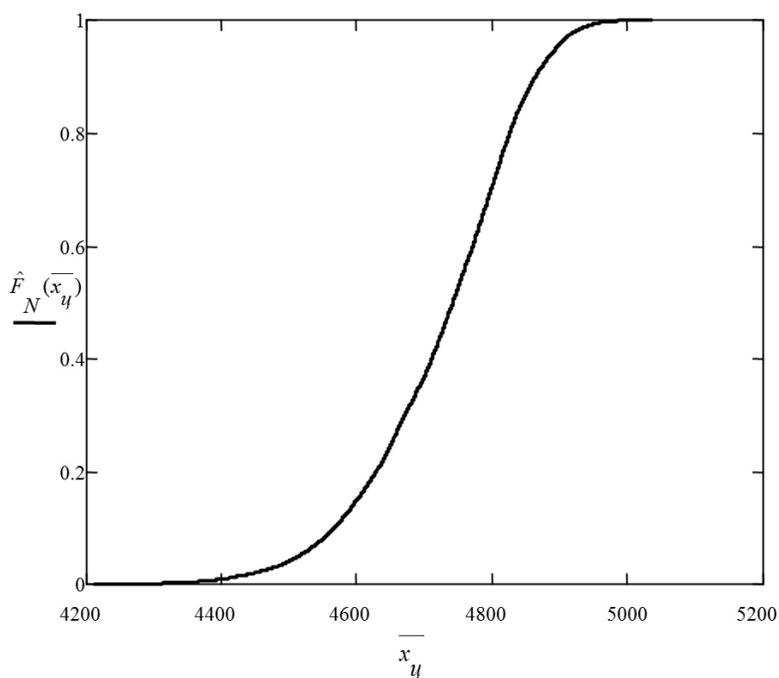


Рис. 5. Графики оценки ф.р. $F_N(\bar{X}_N^u)$ для данных о доходах студентов ТГУ,
 $N = 80$, $T_1 = 2000$, $T_2 = 5000$, $n_2 = 6$, $n_3 = 10$

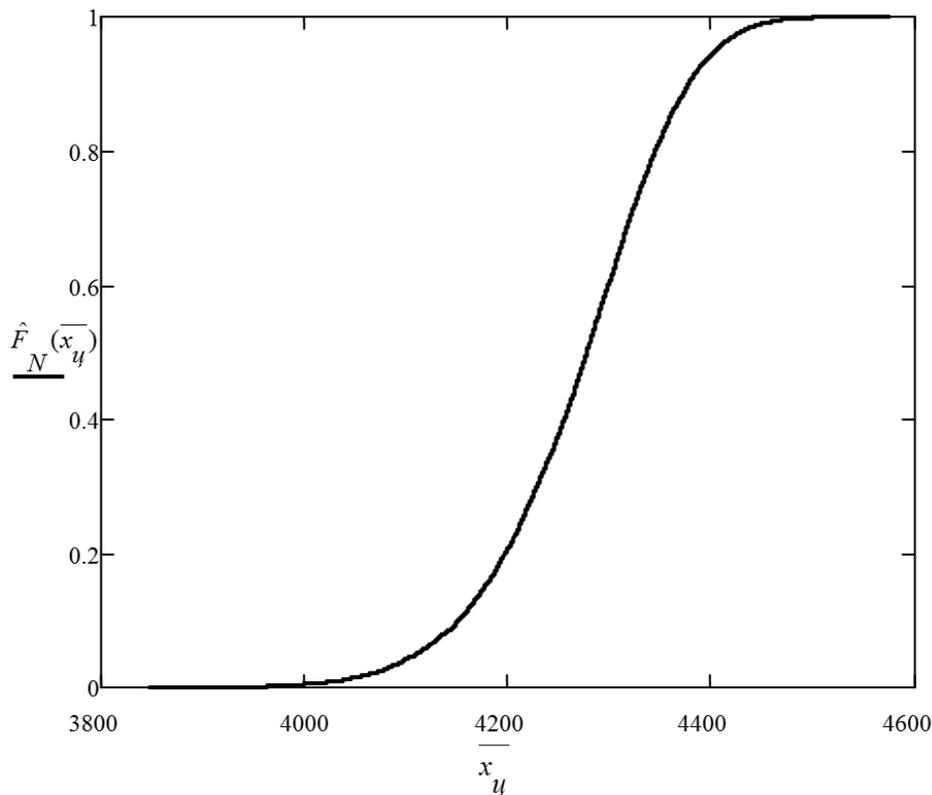


Рис. 6. Графики оценки ф.р. $F_N(\bar{X}_N^u)$ для данных о доходах студентов СибГМУ,

$$N = 80, T_1 = 2000, T_2 = 5000, n_2 = 11, n_3 = 5$$

Полученные распределения позволили получить доверительные интервалы для среднего ежемесячного дохода. Так, для студентов ТГУ, с вероятностью 90% средний ежемесячный доход попадет в интервал [4513.56; 4893.75] руб./мес., а для студентов СИБГМУ — в интервал [4110.86; 4404.26] руб./мес.

Таким образом, можно сделать вывод, что студенты, учащиеся в ТГУ, получают больший доход, чем студенты-медики. Это может быть обусловлено разным уровнем стипендий, выплачиваемых университетами. Или, что наиболее вероятно, число студентов ТГУ, занимающихся подработкой, больше числа подрабатывающих студентов-медиков.

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. Боровков А. А. Математическая статистика. — Новосибирск: Наука; Изд-во Института математики, 1997. 772 с.
3. Эфрон Б. Нетрадиционные методы многомерного статистического анализа: Сб. статей: Пер. с англ./ Предисловие Ю. П. Адлера, Ю. А. Кошевника. — М.: Финансы и статистика, 1988. — 263 с.