АСИМПТОТИЧЕСКОЕ РАСПРЕДЕЛЕНИЕ ПОСЛЕДОВАТЕЛЬНОГО ЭМПИРИЧЕСКОГО ПРОЦЕССА В МОДЕЛИ ПРОПОРЦИОНАЛЬНЫХ ИНТЕНСИВНОСТЕЙ

Х. Ш. Таджиев

Ташкентский государственный экономический университет
Ташкент, Узбекистан
E-mail: asi_96@mail.ru

В модели пропорциональных интенсивностей случайного цензурирования справа доказано свойство асимптотической гауссовости последовательного эмпирического процесса.

Ключевые слова: случайное цензурирование, пропорциональные интенсивности, гауссовские процессы.

Пусть $X_1, X_2, ..., X_n$ – независимые и одинаково распределённые (н.о.р) случайные величины (с.в.) с непрерывной функцией распределения (ф.р.) *F*, подвергаются случайному цензурированию справа другой совокупностью н.о.р. с.в. $Y_1, Y_2, ..., Y_n$ с ф.р. G так, что наблюдается выборка объёма n: $U^{(n)} = \{(Z_1, \delta_1), (Z_2, \delta_2), \ldots, (Z_n, \delta_n)\}$, где $Z_i = \min(X_i, Y_i)$, $\delta_i = I(X_i \le Y_i)$ и $I(\cdot)$ – индикаторная функция. Предположим, что совокупности $\{X_i\}$ и $\{Y_i\}$ независимы. В данной модели случайного цензурирования справа задача состоит в оценивании ф.р. *F* при мешающей ф.р. *G* по выборке $U^{(n)}$, в которой числом наблюдённых X_i является с.в. $v_i = \delta_1 + \ldots + \delta_n$. В данной работе эта задача рассматривается в специальной модели пропорциональных интенсивностей (МПИ), в которой цензурирование является информативным, т.е. существует положительное число β (параметр цензурирования) и имеет место следующее представление ф.р. *G* через *F*:

$$1 - G(t) = (1 - F(t))^{\beta}, \quad t \in R^1. \tag{1}$$

Из (1) следует пропорциональность интегральных функций интенсивностей $\Lambda_F(t) = -\log(1 - F(t))$ и $\Lambda_G(t) = -\log(1 - G(t))$: $\Lambda_G(t) = \beta\Lambda_F(t)$, $t < T_F = \inf\{t \in R^1 : F(t) = 1\}$, что определяет и название модели. Пусть $H(t) = P(Z_i \le t)$. Тогда $H(t) = 1 - (1 - F(t))(1 - G(t))$ и согласно (1):

$$1 - F(t) = (1 - H(t))^p, \quad t \in R^1, \tag{2}$$

где $p = P(\delta_1 = 1) = P(X_i \le Y_i) = \dfrac{1}{1+\beta}$ – вероятность наблюдения X_i, в которой случай β = 0 соответствует отсутствию цензурирования. Авторами [1, 3] независимо друг от друга для *F* была предложена следующая достаточная оценка максимального правдоподобия, получаемая из (2):

$$F_n(t)=1-\left(1-H_n(t)\right)^{p_n}=\begin{cases}0, & t\le Z_{(1)},\\ 1-\left(\dfrac{n-k}{n}\right)^{p_n}, & Z_{(k)}<t\le Z_{(k+1)},\ k=\overline{1,n-1},\\ 1, & t\ge Z_{(n)},\end{cases} \tag{3}$$

Здесь $Z_{(1)}\le Z_{(2)}\le\ldots\le Z_{(n)}$ – вариационный ряд, построенный по совокупности с.в. $Z_1,Z_2,...,Z_n$ а $H_n(t)$ и p_n – эмпирические оценки для $H(t)$ и p соответственно:

$$H_n(t)=\frac{1}{n}\sum_{i=1}^{n}I(Z_i\le t),\quad p_n=\frac{1}{n}\sum_{i=1}^{n}\delta_i. \tag{4}$$

Дальнейшему исследованию и использованию оценки (3) посвящены огромное количество работ (см. подробно обзоры в работах [2, 4–9]). МПИ, называемая также и моделью Козиола-Грина ([4]), характеризуется независимостью с.в. $\{Z_1,...,Z_n\}$ и $\{\delta_1,\ldots,\delta_n\}$. Из этого свойства следует независимость эмпирических оценок $H_n(t)$ и p_n, входящих в формулу (3) и это обеспечивает эффективность многих статистик, использующих оценку F_n по сравнению с другими, основанными на известной оценке Каплана-Мейера (см. [2, 4]).

Анализ имеющейся литературы показал, что последовательный аналог оценки (3) ещё никем не был исследован. Целью данной работы является исследование последовательного эмпирического процесса

$$\left\{Q_n(s;t)=\frac{[ns]}{n^{1/2}}\left(F_{[ns]}(t)-F(t)\right),\quad (s;t)\in[0,1]\times R^1\right\}, \tag{5}$$

где $[a]$ означает целую часть числа a. Заметим, что в МПИ $T_F=F_G=T_H$. Пусть $T<T_F$, $D_T=[0,1]\times[-\infty,T]$ и $D_\infty=[0,1]\times R^1$.

Справедливо утверждение об асимптотической гауссовости случайной последовательности (5) при $n\to\infty$.

Теорема. Последовательность случайных процессов $\{Q_n(s;t),\ (s;t)\in D_T\}$ слабо сходится к центрированному гауссовскому процессу $\{A(s;t),\ (s;t)\in D_T\}$ с ковариационной структурой для $s_1,s_2\in[0,1]$ и $t_1,t_2\in[-\infty,T]$:

$$Cov\left(A(s_1;t_1),A(s_2;t_2)\right)=\min(s_1,s_2)\left(1-F(t_1)\right)\left(1-F(t_2)\right)\cdot$$
$$\cdot\left\{\frac{p^2H\left(\min(t_1,t_2)\right)}{1-H\left(\min(t_1,t_2)\right)}+p(1-p)\log\left(1-H(t_1)\right)\log\left(1-H(t_2)\right)\right\}. \tag{6}$$

Замечание 1. Предельный гауссовский процесс $A(s,t)$ имеет следующее линейное представление

$$A(s;t)=p\left(1-H(t)\right)^{p-1}\cdot B^H(s;t)-C^P(s)\left(1-H(t)\right)^p\log\left(1-H(t)\right),$$

где $\{B^H(s;t),(s;t)\in D_\infty\}$ и $\{C^P(s),\ s\in[0,1]\}$ – независимые центрированные гауссовские процессы, являющиеся предельными для последовательностей эмпирических процессов

$\left\{\frac{[ns]}{n^{1/2}}\left(H_{[ns]}(t)-H(t)\right),(s;t)\in D_{\infty}\right\}$ и $\left\{\frac{[ns]}{n^{1/2}}(p_n-p),\ s\in[0,1]\right\}$ и имеющие следующие ковариационные функции для $s_1,s_2\in[0,1]$ и $t_1,t_2\in R^1$:

$$Cov\left(B^H(s_1;t_1),\ B^H(s_2;t_2)\right)=\min(s_1,s_2)\left[H\left(\min(t_1,t_2)\right)-H(t_1)H(t_2)\right],$$

$$Cov\left(C^p(s_1),\ C^p(s_2)\right)=\min(s_1,s_2)\cdot p(1-p).$$

Замечание 2. При отсутствии цензурирования, p = 1, p_n = 1, Z_i = X_i, $i=\overline{1,n}$, $H(t)\equiv$ $\equiv F(t)$ оценка (3) совпадает с обычной эмпирической оценкой

$$F_n^{э}(t)=\frac{1}{n}\sum_{i=1}^{n}I(X_i\le t),$$

а гауссовский процесс $\{A(s;t),\ (s;t)\in D_{\infty}\}$ согласно (6) имеет ковариацию

$$Cov\left(A(s_1;t_1),\ A(s_1;t_2)\right)=\min(s_1,s_2)\left[F\left(\min(t_1,t_2)-F(t_1)F(t_2)\right)\right],$$

т.е. получаем предельную ковариационную структуру последовательного эмпирического процесса

$$\frac{[ns]}{n^{1/2}}\left(F_{[ns]}^{э}(t)-F(t)\right),\ (s;t)\in D_{\infty}.$$

ЛИТЕРАТУРА

1. *Абдушукуров, А.* О некоторых оценках функции распределения при случайном цензурировании. / А. А. Абдушукуров // Матер. конф. молодых учёных. Инст.мат. АНРУз. Ташкент, 1984. ВИНИТИ. № 8756-В.

2. *Abdushukurov, A.* Статистика неполных наблюдений. / А. А. Абдушукуров. Ташкент. Университет, 2009. 269 с.

3. *Cheng, P.* Maximum likelihood estimation of a survival function under the Koziol-Green proportional hazards model. / P. E. Cheng, G. D. Lin. // Technical Report. B-84-5. 1984. Institute of Statistics, Academica Sinica. Taipei, Taiwan.

4. *Csörgő, S.* Estimation in proportional hazards model of random censorship. / S. Csörgő // Statistics. 1988. V. 19. P. 437–463.

5. *Csörgő, S.* The paradoxical nature of the proportional hazards model of random censorship. / S. Csörgő // Statistics. 1998. V. 31. P. 67–78.

6. *Csörgő, S.* Testing for the partial proportional hazards model of random censorship. / S. Csörgő // Technical Report N 315. 1998. Dpt. Statistics. Univ. Michigan.

7. *Ghorai, J.* The asymptotic distribution of the suprema of the standardized empirical processes under the Koziol-Green model. / J. K. Ghorai. // Statistics Probab. Lett. 1999. V. 41. P. 303–313.

8. *Powlitschko, J.* A comparison of survival function estimators in the Koziol-Green model. / J. Powlitschko // Statistics. 1999. V. 32. P. 277–291.

9. *de Una-Álvares, J.* Kernel distribution function estimation under the Koziol-Green model. / J. de Una-Álvares, W. Gonzáles-Manteiga, C.Cadarso-Suares. // J. Stat. Plan. Infer. 2000. V. 87. P. 199–219.