М. С. Тихов (ННГУ им. Н.И. Лобачевского, Нижний Новгород, Россия).
Оценивание распределений по выборкам случайного объема.

Распределение классической статистики Колмогорова для проверки простой гипотезы согласия с непрерывной функцией распределения (ф.р.) (см. [1] , с.134-141) основано на уклонении эмпирического распределения (ф.р.) $F_n(x)$, построенного по повторной (полной) выборке *фиксированного объема* n. В докладе рассматривается задача оценивания ф. р. $F(x)$ и ее асимптотическое поведение по *неполной* выборке ***случайного объема*** ν_p. Предполагается, что этот объем имеет отрицательное биномиальное распределение $\boldsymbol{NBD}(p, m)$:
$\mathbf{P}(\nu_p = k) = \binom{k-1}{m-1} p^m (1-p)^{k-m}$, $k \in \{m, m+1, ...\}$, $0 < p < 1$, и ν_p не зависит от $\mathcal{U}^{(n)}$.

Пусть $X_1, X_2, \ldots, X_n$ – латентные, независимые и одинаково распределенные случайные величины (н.о.р. с.в.), отрезок $[0, 1]$ – носитель этого распределения, $u_i = i/n, i = 0, 1, \ldots, n$ – точки деления интервала $[0, 1]$ и $W_i = I(X_i < u_i)$ есть индикатор события $(X_i < u_i)$. Требуется построить оценку для ф. р. $F(x)$ и её квантилей $F^{-1}(\alpha), 0 < \alpha < 1$, по выборке $\mathcal{U}^{(n)} = \{(u_i, W_i), i = 0, .., n\}$. Если объем выборки n *фиксирован*, то в качестве оценки $F(x)$ можно взять ядерную оценку типа оценки Надарая-Ватсона $\hat{F}_n(x)$ (см.[2], [3]). Показано, что в этом случае асимптотическое при $n \to \infty$ распределение нормированной разности $\sqrt{nh}\,(\hat{F}_n(x) - F(x))$, где h – ширина окна просмотра данных, является нормальным.

Теорема. *Если* $\hat{F}_p(x) \equiv \hat{F}_{\nu_p}(x)$ *и* $F(x)$ *имеет третью ограниченную производную, то*

$$\sqrt{p}\,(\hat{F}_p(x) - F(x)) \xrightarrow[p \to 0]{d} \zeta_m,$$

где $\zeta_m = \xi_1 + \xi_2 + ... + \xi_m$, $\xi_i, i = 1, ..., m$ *– н.о.р. с.в. с распределением Лапласа.*

Величина ζ_m имеет более тяжелые хвосты, чем у нормального распределения, поэтому доверительный интервал для $F(x)$ будет более широким, нежели для выборок фиксированного объема.

СПИСОК ЛИТЕРАТУРЫ

[1] Колмогоров А.Н. *Об эмпирическом определении закона распределения* (1933). – В книге: Теория вероятностей и математическая статистика. – М.: Наука, 1986 – 535 с.
[2] Tikhov M.S. Statistical Estimation based on Interval Censored Data. -– In: *Param. and Semiparam. Models with Appl. to Rel., Surv. Analisys, and Qual. of Life*: Springer-Verlag: Theor.& Meth., **XLIV** (2004), pp. 209-215.
[3] Okumura H., Naito K. Non-parametric kernel regression for multinomial data. *Journal Multivariate Analysis*, **97** (2006), pp. 2009-2022.