Записки научных
семинаров ПОМИ
Том 328, 2005 г.

Л. В. Гадасина

ОЦЕНКИ ТИПА БЕРРИ-ЭССЕЕНА ДЛЯ МНОГОВЫБОРОЧНЫХ U-СТАТИСТИК

1 Введение

Пусть $X_{j1},\dots,X_{jn_j}$, $j=1,\dots,c$, – независимые выборки случайных величин, принимающих значения в измеримых пространствах $(\mathcal{X}_j,\mathcal{B}_j)$, $j=1,\dots,c$, причем случайные величины внутри каждой выборки не обязательно имеют одинаковие распределения. Предположим, что $n_j\ge m_j\ge 2$, $j=1,\dots,c$. Пусть $\Phi_{I_{\bar m}}:\mathcal{X}_1^{m_1}\times\mathcal{X}_2^{m_2}\times\cdots\times\mathcal{X}_c^{m_c}\to\mathcal{R}$ с $\bar m=(m_1,\dots,m_c)$ и $I_{\bar m}=(I_{m_1},\dots I_{m_c})$, $I_{m_j}=i_{j1},\dots,i_{jm_j}$, $j=1,\dots,c$, – симметричные ядра, такие что $E|\Phi_{I_{\bar m}}(X_{jI_{m_j}})|<\infty$, где $X_{jI_{m_j}}=X_{ji_{j_1}},\dots,X_{ji_{jm_j}}$ при $1\le i_{j1}<\cdots<i_{jm_j}\le n_j$, $j=1,\dots,c$. Рассмотрим многовыборочную U-статистику

$$U_{\bar n}=U_{\bar n}(\Phi_{I_{\bar m}})=\sum\Phi_{I_{\bar m}}(X_{jI_{m_j}},\ I_{m_j}=i_{j1},\dots,i_{jm_j},\ j=1,\dots,c),$$

где $\bar n=(n_1,\dots,n_c)$, и суммирование ведется по всем $1\le i_{j1}<\cdots<i_{jm_j}\le n_j$, $j=1,\dots,c$.

Пусть $\bar d=(d_1,\dots,d_c)$. Определим канонические функции

$$g_{I_{\bar d}}=g_{I_{\bar d}}(X_{jI_{d_j}},I_{d_j}=i_{j1},\dots,i_{jd_j},j=1,\dots,c)=$$

$$=\int\limits_{\mathcal{X}_1}\int\limits_{\mathcal{X}_1}\cdots\int\limits_{\mathcal{X}_c}U_{\bar n}\prod_{j=1}^{c}\prod_{s_j=1}^{d_j}(\delta_{X_{ji_{s_j}}}(dy_{ji_{s_j}})-P_{ji_{s_j}}(dy_{ji_{s_j}}))\prod_{s_j\in I_{m_j}\setminus I_{d_j}}P_{js_j}(dy_{js_j})$$

$$=\sum_{\bar r=\bar 0}^{\bar d}\sum_{\mathcal{L}_{\bar r}}(-1)^{d_1+\cdots+d_c-r_1-\cdots-r_c}\left(E[U_{\bar n}|X_{jL_{r_j}}]-EU_{\bar n}\right),$$

где $\mathcal{L}_{\bar r}=\{1\le l_{j1}<\cdots<l_{jr_j}\le n_j,j=1,\dots,c\}$, $L_{r_j}=l_{j1},\dots,l_{jr_j}\in I_{d_j}$, $j=1,\dots,c$. Тогда в соответствии с каноническим представлением Гефдинга

$$U_{\bar n}=\sum_{\bar d=\bar 0}^{\bar m}\sum_{\mathcal{I}_{\bar d}}g_{I_{\bar d}}=\sum_{\bar d=\bar 0}^{\bar m}U_{\bar n}(g_{I_{\bar d}}),$$

где $\mathcal{I}_{\bar{d}} = \{1 \le i_{j1} < \cdots < i_{jd_j} \le n_j,\ j = 1, \ldots, c\}$ и при $\bar{d} = \bar{0} = (0, \ldots, 0)$ $g_0 = EU_{\bar{n}}$.

Пусть $\sigma^2_{n_j} = \sum\limits_{i_j=1}^{n_j} Eg^2_{i_j}$ и предположим, что $\sigma^2_{\bar{n}} = \sum\limits_{j=1}^{c} \sigma^2_{n_j} > 0$.

Обозначим

$$\Delta_{\bar{n}}(x) = |P(\sigma^{-1}_{\bar{n}}(U_{\bar{n}} - EU_{\bar{n}}) < x) - \Phi(x)|,$$

$$\Delta_{\bar{n}} = \sup_x \Delta_{\bar{n}}(x),$$

где $\Phi(x)$ – функция распределения стандартной нормальной случайной величины.

Получим верхние оценки для величины $\Delta_{\bar{n}}$ и, кроме того, получим условия, при которых величина $\Delta_{\bar{n}} \to 0$ при $\min\{n_1, \ldots, n_c\} \to \infty$, т.е. центральную предельную теорему для многовыборочных U-статистик. Автору неизвестны работы, в которых получена центральная предельная теорема для случая выборок из неодинаково распределенных случайных величин, для случая выборок из одинаково распределенных случайных величин эта теорема имеется в работах [2, 4, 6] и др.

2 Результаты

Обозначим $M_+ = m_1 + \cdots + m_c$, $D_+ = d_1 + \cdots + d_c$. Определим усеченные канонические функции

$$\widetilde{g}_{I_{\bar{d}}} = g_{I_{\bar{d}}} I(|g_{I_{\bar{d}}}| \le \sigma_{\bar{n}}).$$

Теорема 1. *Если* $\sigma^2_{\bar{n}} > 0$, *то для всякого* $1 \le p < 2$ *и* $n_j \ge m_j \ge 2$, $j = 1, \ldots, c$

$$\Delta_{\bar{n}} \le A_{\bar{m}}(\sigma^{-2}_{\bar{n}} \sum_{j=1}^{c} \sum_{i_j=1}^{n_j} Eg^2_{i_j} I(|g_{i_j}| > \sigma_{\bar{n}}) +$$

$$+ \sigma^{-3}_{\bar{n}} \sum_{j=1}^{c} \sum_{i_j=1}^{n_j} E|g_{i_j}|^3 I(|g_{i_j}| \le \sigma_{\bar{n}}) +$$

$$+ \sigma^{-1}_{\bar{n}} \sum_{\bar{d}:D_+ \ge 2} \sum_{\mathcal{I}_{\bar{d}}} E|g_{I_{\bar{d}}}| I(|g_{I_{\bar{d}}}| > \sigma_{\bar{n}}) +$$

$$+\sum_{\bar d:D_+\ge 2}\sigma_{\bar n}^{-D_+-1}\sum_{\mathcal{I}_{\bar d}}E|\prod_{j=1}^{c}\prod_{s_j=1}^{d_j}\widetilde{g}_{i_{s_j}}\widetilde{g}_{I_{\bar d}}|+$$

$$+\frac{1}{(2-p)^3}\sigma_{\bar n}^{-p}\sum_{\bar d:D_+\ge 2}\sum_{\mathcal{I}_{\bar d}}E|g_{I_{\bar d}}|^p I(|g_{I_{\bar d}}|\le\sigma_{\bar n})).$$

Следствие 1. *Пусть* $p\ge 4/3$, $\gamma_{D_+}=\frac{2D_+}{2D_+-1}$ *при* $D_+=1,\dots,M_+$,

$$E|g_{I_{\bar d}}|^{\gamma_{D_+}}<\infty,\qquad \bar d=1,\dots,\bar m,$$

и

$$\Lambda_{D_+,\bar n}(\epsilon)=\sigma_n^{-\gamma_{D_+}}\sum_{\mathcal{I}_{\bar d}}E|g_{I_{\bar d}}|^{\gamma_{D_+}}I(|g_{I_{\bar d}}|>\epsilon\sigma_{\bar n})\to 0$$

для всякого $\epsilon>0$ *и* $\min\{n_1,\dots,n_c\}\to\infty$.

Тогда

$$\Delta_{\bar n}\to 0$$

при $\min\{n_1,\dots,n_c\}\to\infty$.

Следствие 2. *Пусть* $\beta_{D_+}=\frac{2D_++1}{2D_+-1}$ *при* $D_+=1,\dots,M_+$. *Тогда*

$$\Delta_{\bar n}\le A_{\bar m}\sum_{\bar d:D_+\ge 1}\sigma_{\bar n}^{-\beta_{D_+}}\sum_{\mathcal{I}_{\bar d}}E|g_{I_{\bar d}}|^{\beta_{D_+}}.$$

3 Доказательства

Определим канонические функции для многовыборочной U-статистики с ядром $\widetilde{g}_{I_{\bar d}}$:

$$\overline{g}_{I_{\bar d}}=\overline{g}_{I_{\bar d}}(X_{jI_{d_j}},j=1,\dots,c)=$$

$$=\int_{\mathcal{X}_1}\int_{\mathcal{X}_1}\cdots\int_{\mathcal{X}_c}U_{\bar n}(\widetilde{g}_{I_{\bar d}})\prod_{j=1}^{c}\prod_{s_j=1}^{d_j}(\delta_{X_{ji_{s_j}}}(dy_{ji_{s_j}})-P_{ji_{s_j}}(dy_{ji_{s_j}})),$$

и, обозначив $\bar b=(b_1,\dots,b_c)$, где $b_j=1,\dots,d_j$, $j=1,\dots,c$, определим для $\bar b<\bar d=\{b_j:\,b_j<d_j\}_{j=1}^c$

$$\overline{g}_{I_{\bar d}L_{\bar b}}=\overline{g}_{I_{\bar d}L_{\bar b}}(X_{jL_{b_j}},\ L_{b_j}=i_{j1},\dots,i_{jb_j},j=1,\dots,c)=$$

$$= \int\limits_{\mathcal{X}_1}\int\limits_{\mathcal{X}_1}\cdots\int\limits_{\mathcal{X}_c} U_{\bar n}(\widetilde g_{I_{\bar d}})\prod_{j=1}^{c}\prod_{s_j=1}^{b_j}(\delta_{X_{jl_{s_j}}}(dy_{jl_{s_j}})-P_{jl_{s_j}}(dy_{jl_{s_j}}))\prod_{s_j\in I_{d_j}\setminus L_{b_j}} P_{js_j}(dy_{js_j})$$

$$= -\int\limits_{\mathcal{X}_1}\int\limits_{\mathcal{X}_1}\cdots\int\limits_{\mathcal{X}_c} U_{\bar n}(g_{I_{\bar d}}I(|g_{I_{\bar d}}|>\sigma_{\bar n}))\prod_{j=1}^{c}\prod_{s_j=1}^{b_j}(\delta_{X_{jl_{s_j}}}(dy_{jl_{s_j}})-P_{jl_{s_j}}(dy_{jl_{s_j}}))$$

$$\times\prod_{s_j\in I_{d_j}\setminus L_{b_j}} P_{js_j}(dy_{js_j}).$$

Тогда по представлению Гефдинга

$$U_{\bar n}(\widetilde g_{I_{\bar d}}) = \sum_{\bar b=\bar 0}^{\bar m}\sum_{\mathcal{L}_{\bar b}}\overline{g}_{I_{\bar d}L_{\bar b}}(X_{jL_{b_j}},\ j=1,\dots,c)$$

с $\mathcal{L}_{\bar b} = \{1\le l_{j1}<\cdots<l_{jb_j}\le n_j,\ j=1,\dots,c\}$, где $L_{b_j}\in I_{d_j}$, $j=1,\dots,c$.

И если $B_+ = b_1+\cdots+b_c$, то

$$\sum_{\bar d=\bar 0}^{\bar m} U_{\bar n}(\widetilde g_{I_{\bar d}}) = \sum_{\bar d=\bar 0}^{\bar m} EU_{\bar n}(\widetilde g_{I_{\bar d}})+$$

$$+\sum_{\bar d:D_+\ge 1}\sum_{\bar b:B_+\ge 1}\sum_{\mathcal{L}_{\bar b}}\overline{g}_{I_{\bar d}L_{\bar b}}(X_{jL_{b_j}},\ j=1,\dots,c) =$$

$$= -\sum_{\bar d:D_+\ge 1}\sum_{\mathcal{I}_{\bar d}} E|g_{I_{\bar d}}|I(|g_{I_{\bar d}}|>\sigma_{\bar n})+$$

$$+\sum_{\bar d:D_+\ge 1}\sum_{j=1}^{c}\sum_{l_j=1}^{n_j}\overline{g}_{I_{\bar d}l_j}(X_{jl_j})+$$

$$+\sum_{\bar d:D_+\ge 2}\sum_{\mathcal{I}_{\bar d}}\overline{g}_{I_{\bar d}}(X_{jI_{d_j}},\ j=1,\dots,c)+$$

$$+\sum_{\bar b:B_+\ge 2}\sum_{\bar d:D_+>B_+}\sum_{\mathcal{L}_{\bar b}}\overline{g}_{I_{\bar d}L_{\bar b}}(X_{jL_{b_j}},\ j=1,\dots,c).$$

Пусть

$$\theta = -\sum_{\bar d:D_+\ge 1}\sum_{\mathcal{I}_{\bar d}} E|g_{I_{\bar d}}|I(|g_{I_{\bar d}}|>\sigma_{\bar n}),$$

$$h_{\bar d l_j} = \overline{g}_{I_{\bar d} l_j}(X_{j l_j}), \quad \bar d : D_+ \geq 2,$$

$$\tilde h_{\bar d l_j} = h_{\bar d l_j} I(|h_{\bar d l_j}| \leq \sigma_{\bar n}),$$

$$\bar h_{\bar d l_j} = \tilde h_{\bar d l_j} - E\tilde h_{\bar d l_j}, \quad j = 1, \ldots, c,$$

$$\bar\sigma^2_{n_j} = \sum_{l_j=1}^{n_j} (\overline{g}_{l_j} + \sum_{\bar d : D_+ \geq 2} \bar h_{\bar d l_j})^2, \quad \bar\sigma^2_{\bar n} = \sum_{j=1}^{c} \bar\sigma^2_{n_j},$$

$$z_{j i_j} = \bar\sigma_{\bar n}^{-1}(\overline{g}_{l_j} + \sum_{\bar d : D_+ \geq 2} \bar h_{\bar d l_j}), \quad \hat z_{j i_j} = z_{j i_j} + \bar\sigma_{\bar n}^{-1} E(\tilde g_{l_j} + \sum_{\bar d : D_+ \geq 2} \tilde h_{\bar d l_j}),$$

$$Z_{n_j} = \sum_{i_j=1}^{n_j} z_{j i_j}, \quad Z_{\bar n} = \sum_{j=1}^{c} Z_{n_j} = \sum_{j=1}^{c} \sum_{i_j=1}^{n_j} z_{j i_j}.$$

Применив неравенство о срезках, получим

$$\Delta_{\bar n}(x) \leq |P(\sigma_{\bar n}^{-1}(\sum_{\bar d=\bar 0}^{\bar m} U_{\bar n}(\tilde g_{I_{\bar d}}) - EU_{\bar n}) < x) - \Phi(x)| + \sum_{\bar d : D_+ \geq 1} \sum_{\mathcal{I}_{\bar d}} P(|g_{I_{\bar d}}| > \sigma_{\bar n})$$

$$\leq |P(\sigma_{\bar n}^{-1}(\theta + \sum_{j=1}^{c} \sum_{l_j=1}^{n_j} (\overline{g}_{l_j} + \sum_{\bar d : D_+ \geq 2} h_{\bar d l_j}) + \bar\sigma_{\bar n} W_{\bar n} + \bar\sigma_{\bar n} V_{\bar n} < x) - \Phi(x)| +$$

$$+ \sum_{\bar d : D_+ \geq 1} \sum_{\mathcal{I}_{\bar d}} P(|g_{I_{\bar d}}| > \sigma_{\bar n}),$$

где

$$V_{\bar n} = \bar\sigma_{\bar n}^{-1} \sum_{\bar b : B_+ \geq 2} \sum_{\bar d : D_+ > B_+} \sum_{\mathcal{L}_{\bar b}} \overline{g}_{I_{\bar d} L_{\bar b}}, \qquad W_{\bar n} = \bar\sigma_{\bar n}^{-1} \sum_{\bar d : D_+ \geq 2} \sum_{\mathcal{I}_{\bar d}} \overline{g}_{I_{\bar d}}.$$

По неравенству о срезках для $h_{\bar d l_j}$, получим

$$\Delta_{\bar n}(x) \leq$$

$$|P(\sigma_{\bar n}^{-1}(\theta + \bar\sigma_{\bar n} Z_{\bar n} + \bar\sigma_{\bar n} W_{\bar n} + \bar\sigma_{\bar n} V_{\bar n} + \sum_{j=1}^{c} \sum_{l_j=1}^{n_j} \sum_{\bar d : D_+ \geq 2} E\tilde h_{\bar d l_j}) < x) - \Phi(x)| +$$

$$+ \sum_{\bar d : D_+ \geq 1} \sum_{\mathcal{I}_{\bar d}} P(|g_{I_{\bar d}}| > \sigma_{\bar n}) + \sum_{j=1}^{c} \sum_{l_j=1}^{n_j} \sum_{\bar d : D_+ \geq 2} P(|h_{\bar d l_j}| > \sigma_{\bar n}) =$$

$$= |P(\bar\sigma_{\bar n}\sigma_{\bar n}^{-1}(Z_{\bar n}+V_{\bar n}+W_{\bar n})+\varphi_{\bar n}<x)-\Phi(x)|+\psi_{\bar n},$$

где

$$\varphi_{\bar n}=\sigma_{\bar n}^{-1}(\theta+\sum_{j=1}^{c}\sum_{l_j=1}^{n_j}\sum_{D_+\geq 2}E\tilde h_{\bar d l_j}),$$

$$\psi_{\bar n}=\sum_{D_+\geq 1}\sum_{\mathcal{I}_{\bar d}}P(|g_{I_{\bar d}}|>\sigma_{\bar n})+\sum_{j=1}^{c}\sum_{l_j=1}^{n_j}\sum_{D_+\geq 2}P(|h_{\bar d l_j}|>\sigma_{\bar n}).$$

Тогда

$$\Delta_{\bar n}\leq\sup_x|P(Z_{\bar n}+V_{\bar n}+W_{\bar n}<x)-\Phi(x)|+$$

$$+\sup_x|\Phi(x+\varphi_{\bar n})-\Phi(x)|+\sup_x|\Phi(\frac{\bar\sigma_{\bar n}}{\sigma_{\bar n}}x)-\Phi(x)|+\psi_{\bar n},$$

где по свойствам функции Φ

$$\sup_x|\Phi(x+\varphi_{\bar n})-\Phi(x)|\leq\frac{1}{\sqrt{2\pi}}|\varphi_{\bar n}|, \tag{1}$$

$$\sup_x|\Phi(\frac{\bar\sigma_{\bar n}}{\sigma_{\bar n}}x)-\Phi(x)|\leq\frac{1}{\sqrt{2\pi e}}|\frac{\bar\sigma_{\bar n}^2}{\sigma_{\bar n}^2}-1|\frac{\sigma_{\bar n}}{\bar\sigma_{\bar n}}. \tag{2}$$

Пусть

$$\delta=\Big(\sum_{j=1}^{c}\sum_{i_j=1}^{n_j}E|z_{j\,i_j}|\Big)^{-1},$$

не уменьшая общности, можно считать, что $\delta>2^{3+M_+}$, тогда по неравенству Эссеена первое слагаемое в $\Delta_{\overline{n}}$ не превосходит

$$\frac{2}{\pi}\int_0^{\delta}|E\exp(itZ_{\bar n})-\exp(-t^2/2)|t^{-1}dt+$$

$$+\frac{2}{\pi}\int_0^{2^{3+M_+}}|E\exp(it(Z_{\bar n}+V_{\bar n}+W_{\bar n}))-E\exp(itZ_{\bar n})|t^{-1}dt+$$

$$+\frac{2}{\pi}\int_{2^{3+M_+}}^{\delta}|E\exp(it(Z_{\bar n}+V_{\bar n}+W_{\bar n}))-E\exp(itZ_{\bar n})|t^{-1}dt+\frac{24}{\pi\sqrt{2\pi}}\delta^{-1}=$$

$$= \varepsilon_1 + \varepsilon_2 + \varepsilon_3 + \frac{24}{\pi\sqrt{2\pi}}\delta^{-1}.$$

Легко увидеть, что, если $M_\times = m_1 \cdot \ldots \cdot m_c$,

$$\varepsilon_1 \le \frac{16}{\pi}\bar{\sigma}_{\bar{n}}^{-3}\sum_{j=1}^{c}\sum_{i_j=1}^{n_j} E|\tilde{g}_{i_j}|^3 +$$

$$+\frac{32}{\pi}\bar{\sigma}_{\bar{n}}^{-3}\sigma_{\bar{n}}^2(M_\times - c)^2 M_\times \sum_{D_+\ge 2}\sum_{\mathcal{I}_{\bar{d}}} E|g_{I_{\bar{d}}}|I(|g_{I_{\bar{d}}}| > \sigma_{\bar{n}}).$$

Для слагаемого с δ^{-1} оценка аналогична.

Утверждение 1. *Пусть* $\xi_{j1}^{(k_j)}, \ldots, \xi_{jn_i}^{(k_j)}$, $k_j = 1, \ldots, d_j$, $j = 1, \ldots, c$, – *последовательности выборок, состоящих из независимых одинаково распределенных случайных величин, таких что* $\xi_{ji_1}^{(1)}, \ldots, \xi_{ji_{d_j}}^{(d_j)}$ *независимы попарно различных вторых индексах, и, кроме того, эти последовательности не зависят от* $X_{j1}, \ldots, X_{jn_j}$, $j = 1, \ldots, c$. *Тогда*

$$E|\sum_{\mathcal{I}_{\bar{d}}} \bar{g}_{I_{\bar{d}}}|^p \le 2^{(2-p)D_+}\sum_{\mathcal{I}_{\bar{d}}} E|\bar{g}_{I_{\bar{d}}}|^p, \quad 1 \le p \le 2,$$

и

$$E|\sum_{\mathcal{I}_{\bar{d}}}\prod_{j=1}^{c}\xi_{ji_1}^{(1)} \cdot \ldots \cdot \xi_{ji_{d_j}}^{(d_j)}\bar{g}_{I_{\bar{d}}}|^p \le$$

$$\le 2^{(2-p)D_+}\prod_{j=1}^{c}\prod_{s_j=1}^{d_j} E|\xi_{j1}^{(s_j)}|^p \sum_{\mathcal{I}_{\bar{d}}} E|\bar{g}_{I_{\bar{d}}}|^p, \quad 1 \le p \le 2.$$

Оценка ε_2**.** Нам понадобится следующая лемма.

Лемма 1. *При всех* $0 \le t \le \delta$, $1 \le i_{j1} < \cdots < i_{jd_j} \le n_j$ *с* $d_j = 1, \ldots, m_j$, $j = 1, \ldots, c$

$$\prod_{j=1}^{c}\prod_{k_j \ne I_{d_j}} |E\exp(itz_{jk_j})| \le \exp(-t^2(1/3 - D_+ 2^{-(2M+9)/3})).$$

Доказательство. Неравенство получается непосредственным обобщением леммы 1 для одновыборочных U-статистик из [3].

Представим подинтегральное выражение ε_2 следующим образом:

$$E\exp(it(Z_{\bar n}+V_{\bar n}+W_{\bar n}))-E\exp(itZ_{\bar n})=$$

$$=E\exp(it(Z_{\bar n}+V_{\bar n}))(\exp(itW_{\bar n})-1-itW_{\bar n})+itE\exp(itZ_{\bar n})W_{\bar n}+$$

$$+itE\exp(itZ_{\bar n})(\exp(itV_{\bar n})-1)W_{\bar n}+itE\exp(itZ_{\bar n})(\exp(itV_{\bar n})-1)=$$

$$=L_1+L_2+L_3+L_4.$$

Оценки для каждого слагаемого получаются таким же образом, как и для случая одновыборочных U-статистик, см. [3]. Запишем подробно только для L_2. Поскольку канонические функции $\bar g_{I_{\bar d}}$ обладают свойством полной вырожденности, имеем

$$|L_2|\le|t||E\exp(itZ_{\bar n})W_{\bar n}|=$$

$$=|t|\bar\sigma_{\bar n}^{-1}\sum_{D_+\ge 2}\sum_{\mathcal{I}_{\bar d}}|E\exp(it\sum_{j=1}^{c}\sum_{k_j\ne I_{a_j}}z_{jk_j})|\times$$

$$\times|E\prod_{j=1}^{c}\prod_{k_j=1}^{d_j}(\exp(it\hat z_{ji_{k_j}})-1)\bar g_{I_{\bar d}}|.$$

Тогда по лемме 1,

$$|L_2|\le|t|^{M_++1}\exp(-t^2(1/3-M_+2^{-(2M+9)/3}))\times$$

$$\times\bar\sigma_{\bar n}^{-1}\sum_{D_+\ge 2}\sum_{\mathcal{I}_{\bar d}}E|\prod_{j=1}^{c}\prod_{k_j=1}^{d_j}\hat z_{ji_{k_j}}\bar g_{I_{\bar d}}|.$$

Оценка ε_3**.** Используем метод рандомизации. Зафиксируем $t\in[2^{3+M_+},\delta]$ и рассмотрим $n_1+\cdots+n_c$ независимых одинаково распределенных случайных величин $\alpha_{j1},\dots,\alpha_{jn_j}$, $j=1,\dots,c$, таких что они не зависят от $X_{j1},\dots,X_{jn_j}$, $j=1,\dots,c$, и имеют распределение Бернулли

$$P(\alpha_{jk_j}=1)=1-P(\alpha_{jk_j}=0)=f(t)=3\cdot2^{M_+}\cdot t^{-2}\log(t)$$

при всех $k_j=1,\dots,n_j$, $j=1,\dots,c$.

Заметим, что $f(t)\in[0,1]$ при всех t. Тогда можем записать следующее представление:

$$X_{jk_j}\stackrel{d}{=}\alpha_{jk_j}\overline{X}_{jk_j}+(1-\alpha_{jk_j})\widetilde{X}_{jk_j},\quad k_j=1,\dots,n_j,\quad j=1,\dots,c,$$

где $\overline{X}_{jk_j}, \widetilde{X}_{jk_j}$ – независимые копии X_{jk_j} при всех $k_j = 1, \dots, n_j$, $j = 1, \dots, c$. Тогда

$$Z_{n_j} \stackrel{d}{=} \sum_{i_j=1}^{n_j} \alpha_{ji_j}\bar{z}_{ji_j} + \sum_{i_j=1}^{n_j}(1-\alpha_{ji_j})\tilde{z}_{ji_j} = Z_{n_j}^{(1)}(\overline{X}_j) + Z_{n_j}^{(2)}(\widetilde{X}_j),$$

$$Z_{\bar{n}} = \sum_{j=1}^{c} Z_{n_j}^{(1)} + \sum_{j=1}^{c} Z_{n_j}^{(2)} = Z_{\bar{n}}^{(1)} + Z_{\bar{n}}^{(2)}.$$

Кроме того,

$$W_{\bar{n}} \stackrel{d}{=} \bar{\sigma}_{\bar{n}}^{-1} \sum_{D_+\geq 2} \sum_{\bar{k}=\bar{0}}^{\bar{d}} \sum_{\mathcal{R}_{\bar{k}}} \sum_{\mathcal{I}_{\bar{k}}^{\bar{d}},\ I_{k_j}^{d_j} \neq R_{k_j}} \prod_{j=1}^{c} \prod_{s_j=1}^{k_j} \alpha_{jr_{s_j}} \prod_{s_j=k_j+1}^{d_j} (1-\alpha_{ji_{s_j}}) \times$$

$$\times \bar{g}_{I_{\bar{d}}}(\overline{X}_{jR_{k_j}}, \widetilde{X}_{jI_{k_j}^{d_j}},\ j = 1, \dots, c) = \sum_{D_+\geq 2} \sum_{\bar{k}=\bar{0}}^{\bar{d}} W_{\bar{d}\bar{k}},$$

$$V_{\bar{n}} \stackrel{d}{=} \bar{\sigma}_{\bar{n}}^{-1} \sum_{\bar{b}:B_+\geq 2} \sum_{\bar{d}:D_+>B_+} \sum_{\bar{k}=\bar{0}}^{\bar{b}} \sum_{\mathcal{R}_{\bar{k}}} \sum_{\mathcal{L}_{\bar{k}}^{\bar{b}},\ L_{k_j}^{b_j} \neq R_{k_j}} \prod_{j=1}^{c} \prod_{s_j=1}^{k_j} \alpha_{jr_{s_j}} \prod_{s_j=k_j+1}^{b_j} (1-\alpha_{jl_{s_j}})$$

$$\times \bar{g}_{I_{\bar{d}}L_{\bar{b}}}(\overline{X}_{jR_{k_j}}, \widetilde{X}_{jL_{k_j}^{b_j}},\ j = 1, \dots, c) = \sum_{B_+\geq 2} \sum_{\bar{d}:D_+>B_+} \sum_{\bar{k}=\bar{0}}^{\bar{b}} V_{\bar{d}\bar{b}\bar{k}},$$

где $\mathcal{I}_{\bar{k}}^{\bar{d}} = \{1 \leq i_{jk_j+1} < \dots < i_{jd_j} \leq n_j,\ j = 1, \dots, c\}$, $\mathcal{L}_{\bar{k}}^{\bar{b}} = \{1 \leq l_{jk_j+1} < \dots < l_{jb_j} \leq n_j,\ j = 1, \dots, c\}$.

Лемма 2. *Обозначим $\alpha_{j1}, \dots, \alpha_{jn_j}$, $j = 1, \dots, c$, через α. Тогда для всех $t \in [2^{3+M_+}, \delta]$*

$$E|E[\exp(itZ_{\bar{n}}^{(1)})|\alpha]|^2 \leq t^{-2^{M_+}},$$

$$|E[\exp(itZ_{\bar{n}}^{(2)})|\alpha]|^2 \leq t^{2^{M_+}} e^{-t^2/3},$$

$$E|E[\exp(it\sum_{j=1}^{c}\sum_{s_j=1,s_j\neq I_{d_j}}^{n_j} \alpha_{js_j}z_{js_j})|\alpha]|^2 \leq t^{-2^{M_+}(1-3D_+2^{-2M_+/3-2})},$$

$$E|E[\exp(it\sum_{j=1}^{c}\sum_{s_j\neq I_{d_j}} (1-\alpha_{js_j})z_{js_j})|\alpha]|^2 \leq$$

$$\leq t^{2^{M_+}+(1-3D_+2^{-2M_+/3-2})}\exp(-t^2(1/3-D_+2^{-2M_+/3-2})).$$

Доказательство. Является распространением леммы 1 для одновыборочных U-статистик из [3] на случай многовыборочных.

Рассмотрим следующее разложение подинтегрального выражения в ε_3

$$E\exp(it(Z_{\bar n}+V_{\bar n}+W_{\bar n}))-E\exp(itZ_{\bar n})=$$

$$=E\exp(it(Z_{\bar n}+W_{\bar n}))-E\exp(itZ_{\bar n})+$$

$$+E\exp(it(Z_{\bar n}+V_{\bar n}+W_{\bar n}))-E\exp(it(Z_{\bar n}+W_{\bar n}))=I_1+I_2.$$

Тогда во введенных обозначениях

$$I_1\stackrel{d}{=}E\exp(it(Z_{\bar n}^{(1)}+Z_{\bar n}^{(2)}+\sum_{D_+\geq 2}\sum_{\bar k=\bar 0}^{\bar d}W_{\bar d\bar k}))-E\exp(it(Z_{\bar n}^{(1)}+Z_{\bar n}^{(2)}))=$$

$$=E\exp(it(Z_{\bar n}^{(1)}+Z_{\bar n}^{(2)}+\sum_{D_+\geq 2}W_{\bar d\bar d}))-E\exp(it(Z_{\bar n}^{(1)}+Z_{\bar n}^{(2)}))+$$

$$+E\exp(it(Z_{\bar n}^{(1)}+Z_{\bar n}^{(2)}+\sum_{D_+\geq 2}(W_{\bar d\bar d}+W_{\bar d\bar 0})))-$$

$$-E\exp(it(Z_{\bar n}^{(1)}+Z_{\bar n}^{(2)}+\sum_{D_+\geq 2}W_{\bar d\bar d}))+$$

$$+E\exp(it(Z_{\bar n}^{(1)}+Z_{\bar n}^{(2)}+\sum_{D_+\geq 2}\sum_{\bar k=\bar 0}^{\bar d}W_{\bar d\bar k}))-E\exp(it(Z_{\bar n}^{(1)}+Z_{\bar n}^{(2)}+$$

$$+\sum_{D_+\geq 2}(W_{\bar d\bar d}+W_{\bar d\bar 0})))=I_{11}+I_{12}+I_{13}.$$

Рассмотрим I_{11}:

$$I_{11}=E[\exp(it(Z_{\bar n}^{(1)}+Z_{\bar n}^{(2)}))(\exp(it\sum_{D_+\geq 2}W_{\bar d\bar d})-1-it\sum_{D_+\geq 2}W_{\bar d\bar d})]+$$

$$+itE[\exp(it(Z_{\bar n}^{(1)}+Z_{\bar n}^{(2)}))\sum_{D_+\geq 2}W_{\bar d\bar d}]=I_{111}+I_{112}.$$

Заметим, что $Z_{\bar{n}}^{(2)}$ не зависит от $Z_{\bar{n}}^{(1)}$ и $W_{\bar{d}\bar{d}}$ при фиксированной последовательности α. Тогда по утверждению 1 и лемме 2 получим, аналогично [3],

$$|I_{111}| \le A_{\bar{m}}|t|^{p+2^{M_+-1}} e^{-t^2/6}\overline{\sigma}_{\bar{n}}^{-p} \sum_{D_+\ge 2}\sum_{\mathcal{I}_{\bar{d}}} E|\overline{g}_{I_{\bar{d}}}|^p, \quad 1\le p\le 2,$$

$$|I_{112}| \le$$

$$A_{\bar{m}}|t|^{3M_+2^{M_+/3-3}} e^{-t^2/6}\log^{M_+/2}(t)\overline{\sigma}_{\bar{n}}^{-1} \sum_{D_+\ge 2}\sum_{\mathcal{I}_{\bar{d}}} E|\prod_{j=1}^{c}\prod_{s_j=1}^{d_j}\widehat{z}_{ji_{s_j}}\overline{g}_{I_{\bar{d}}}|.$$

Выражение I_{12} раскладывается на три слагаемых:

$$I_{12} = E[\exp(it(Z_{\bar{n}}^{(1)}+Z_{\bar{n}}^{(2)}))(\exp(it\sum_{D_+\ge 2} W_{\bar{d}\bar{d}})-1)(\exp(it\sum_{D_+\ge 2} W_{\bar{d}0})-1)]$$

$$+E[\exp(it(Z_{\bar{n}}^{(1)}+Z_{\bar{n}}^{(2)}))(\exp(it\sum_{D_+\ge 2} W_{\bar{d}0})-1-it\sum_{D_+\ge 2} W_{\bar{d}0})]+$$

$$+itE[\exp(it(Z_{\bar{n}}^{(1)}+Z_{\bar{n}}^{(2)}))\sum_{D_+\ge 2} W_{\bar{d}0}],$$

при этом, используя тот факт, что при фиксированной последовательности α величина $Z_{\bar{n}}^{(1)}$ не зависит от $Z_{\bar{n}}^{(2)}$ и $W_{\bar{d}0}$, и приемы из [3], получим

$$|I_{12}| \le A_{\bar{m}}\,((|t|^{p-4/p}f(t)^{2/p}+|t|^{p-2^{M_+-1}})\overline{\sigma}_{\bar{n}}^{-p} \sum_{D_+\ge 2}\sum_{\mathcal{I}_{\bar{d}}} E|\overline{g}_{I_{\bar{d}}}|^p +$$

$$+|t|^{1+M_+-3M_+2^{M_+/3-3}}\exp(-\frac{t^2}{2}(1/3-M_+2^{-2M_+/3-2}))\log^{M_+/2}(t)\overline{\sigma}_{\bar{n}}^{-1}\times$$

$$\times\sum_{D_+\ge 2}\sum_{\mathcal{I}_{\bar{d}}} E|\prod_{j=1}^{c}\prod_{s_j=1}^{d_j}\widehat{z}_{ji_{s_j}}\overline{g}_{I_{\bar{d}}}|), \quad 1\le p<2.$$

Рассмотрим I_{13}:

$$I_{13} = E[\exp(it(Z_{\bar{n}}^{(1)}+Z_{\bar{n}}^{(2)}+\sum_{D_+\ge 2} W_{\bar{d}0}))\times$$

$$\times(\exp(it\sum_{D_+\ge 2} W_{\bar{d}\bar{d}})-1)(\exp(it\sum_{D_+\ge 2}\sum_{\bar{k}:1\le K_+<D_+} W_{\bar{d}\bar{k}})-1)]+$$

$$+E[\exp(it(Z_{\bar n}^{(1)}+Z_{\bar n}^{(2)}+\sum_{D_+\ge 2}W_{\bar d 0}))\times$$

$$\times(\exp(it\sum_{D_+\ge 2}\sum_{\bar k:2\le K_+<D_+}W_{\bar d\bar k})-1-it\exp(it\sum_{D_+\ge 2}\sum_{\bar k:2\le K_+<D_+}W_{\bar d\bar k}))]+$$

$$+itE[\exp(it(Z_{\bar n}^{(1)}+Z_{\bar n}^{(2)}+\sum_{D_+\ge 2}W_{\bar d 0}))\sum_{D_+\ge 2}\sum_{\bar k:2\le K_+<D_+}W_{\bar d\bar k}]=$$

$$=I_{131}+I_{132}+I_{133}.$$

При этом по утверждению 1, лемме 2 и соотношению $|e^{ix}-1-ix|\le 2|x|^p$, $1\le p\le 2$,

$$|I_{131}|\le A_{\bar m}|t|^p f(t)^{2/p+1/q}\overline{\sigma}_{\bar n}^{-p}\sum_{D_+\ge 2}\sum_{\mathcal{I}_{\bar d}}E|\overline{g}_{I_{\bar d}}|^p,\quad 1\le p\le 2,$$

$$|I_{132}|\le A_{\bar m}|t|^p f(t)\overline{\sigma}_{\bar n}^{-p}\sum_{D_+\ge 2}\sum_{\mathcal{I}_{\bar d}}E|\overline{g}_{I_{\bar d}}|^p,\quad 1\le p<2.$$

Для оценки I_{133} нам снова понадобится рандомизация. Пусть $\eta_{j1}^{(s_j)},\dots,\eta_{jn_j}^{(s_j)}$, $s_j=1,\dots,m_j$, $j=1,\dots,c$ – последовательность независимых одинаково распредленных случайных величин, независящих от α и $X_{j1},\dots,X_{jn_j}$, $j=1,\dots,c$, с распределением Бернулли

$$P(\eta_{jk_j}^{(s_j)}=1)=1-P(\eta_{jk_j}^{(s_j)}=0)=1/2,$$

при $k_j=1,\dots,n_j$, $s_j=1,\dots,m_j$, $j=1,\dots,c$. Обозначим эту последовательность через η.

Тогда при всех $j=1,\dots,c$

$$\overline{X}_{ji_j}\stackrel{d}{=}\eta_{ji_j}^{(1)}Y_{ji_j}^{(1)}+(1-\eta_{ji_j}^{(1)})\overline{Y}_{ji_j}^{(1)},$$

$$Y_{ji_j}^{(1)}\stackrel{d}{=}\eta_{ji_j}^{(2)}Y_{ji_j}^{(2)}+(1-\eta_{ji_j}^{(2)})\overline{Y}_{ji_j}^{(2)},$$

$$\dots$$

$$Y_{ji_j}^{(m_j)}\stackrel{d}{=}\eta_{ji_j}^{(m_j+1)}Y_{ji_j}^{(m_j+1)}+(1-\eta_{ji_j}^{(m_j+1)})\overline{Y}_{ji_j}^{(m_j+1)},$$

где $Y_{ji_j}^{(1)}$, $\overline{Y}_{ji_j}^{(1)}$ – независимые копии $\overline{X}_{ji_j}$, и $Y_{ji_j}^{(s_j+1)}$, $\overline{Y}_{ji_j}^{(s_j+1)}$ – независимые копии $Y_{ji_j}^{(s_j)}$ при $s_j=1,\dots,m_j$.

Тогда

$$Z_{n_j}^{(1)} + Z_{n_j}^{(2)} = Z_{n_j 0}^{(1)} + Z_{n_j 0}^{(2)} = Z_{n_j k_j}^{(1)}(Y_j^{(k_j)}) + \sum_{r_j=0}^{k_j} Z_{n_j r_j}^{(2)}(\overline{Y}_j^{(r_j)}),$$

где

$$Z_{n_j k_j}^{(1)} = \sum_{i_j=0}^{n_j} \prod_{s_j=1}^{k_j} \eta_{j i_j}^{(s_j)} \alpha_{j i_j} z_{j i_j}(Y_j^{(k_j)}),$$

$$Z_{n_j r_j}^{(2)} = \sum_{i_j=0}^{n_j} (1 - \eta_{j i_j}^{(r_j)}) \prod_{s_j=1}^{r_j} \eta_{j i_j}^{(s_j)} \alpha_{j i_j} \overline{z}_{j i_j}(\overline{Y}_j^{(r_j)}).$$

Рассмотрим разложение для первой выборки $\overline{X}_{1R_{k_1}}$, $R_{k_1} = r_{11}, \ldots, r_{1k_1}$, тогда получим представление (см. [3])

$$W_{\bar{d}\bar{k}} \stackrel{d}{=} \sum_{h_1=0}^{k_1-1} \sum_{1 \le a_{h_1} < a_{h_1-1} < \ldots < k_1} (W_{\bar{d}\bar{k} a_{h_1} a_{h_1}} + W_{\bar{d}\bar{k} a_{h_1} 0}),$$

где $W_{\bar{d}\bar{k} a_{h_1} a_{h_1}}$ и $W_{\bar{d}\bar{k} a_{h_1} 0}$ определяются как и в [3].

Рассмотрев разложение для второй выборки $\overline{X}_{2R_{k_2}}$, получаем

$$W_{\bar{d}\bar{k}} \stackrel{d}{=} \sum_{h_1=0}^{k_1-1} \sum_{1 \le a_{h_1} < a_{h_1-1} < \ldots < k_1} \sum_{h_2=0}^{k_2-1} \sum_{1 \le a_{h_2} < a_{h_2-1} < \ldots < k_2} \times$$

$$\times (W_{\bar{d}\bar{k} a_{h_1} a_{h_1} a_{h_2} a_{h_2}} + W_{\bar{d}\bar{k} a_{h_1} a_{h_1} a_{h_2} 0} + W_{\bar{d}\bar{k} a_{h_1} 0 a_{h_2} a_{h_2}} + W_{\bar{d}\bar{k} a_{h_1} 0 a_{h_2} 0}),$$

и аналогично, беря разложения для всех выборок, получаем

$$W_{\bar{d}\bar{k}} \stackrel{d}{=} \sum_{h_j=0}^{k_j-1} \sum_{1 \le a_{h_j} < a_{h_j-1} < \ldots < k_j} \sum_{\bar{A}} W_{\bar{d}\bar{k}\bar{A}},$$

где $j = 1, \ldots, c$ и $\bar{A}$ – комбинация из $a_{h_j}0$ и $a_{h_j} a_{h_j}$, $j = 1, \ldots, c$. Таких комбинаций 2^c, и

$$W_{\bar{d}\bar{k}\bar{A}} = \sum_{\mathcal{L}_{\bar{a}_h}} \sum_{\mathcal{L}_{\bar{a}_h}^{\bar{a}_h - 1}} \cdots \sum_{\mathcal{R}_{\bar{a}_1}^{\bar{k}}} \sum_{\mathcal{I}_{\bar{k}}^{\bar{d}}} \prod_{j=1}^{c} \prod_{s_j=1}^{a_{1_j}} \alpha_{j l_{s_j}} \prod_{s_j=a_{1_j}+1}^{k_j} \alpha_{j r_{s_j}} \prod_{s_j=k_j+1}^{d_j} (1 - \alpha_{j i_{s_j}})$$

$$\times \prod_{s_j=a_{1_j}+1}^{k_j} (1 - \eta_{j r_{s_j}}^{(1)}) \times \cdots \times \prod_{s_j=a_{h_j}+1}^{a_{h_j-1}} \Big(\prod_{b_j=1}^{h_j-1} \eta_{j l_{s_j}}^{(b_j)} (1 - \eta_{j l_{s_j}}^{(h_j)}) \Big) \times$$

$$\times \prod_{j\in J_1}\Big(\prod_{s_j=1}^{a_{h_j}}\prod_{b_j=1}^{h_j+1}\eta_{jl_{s_j}}^{(b_j)}\Big)\prod_{j\in J_2}\Big(\prod_{s_j=1}^{a_{h_j}}\prod_{b_j=1}^{h_j}\eta_{jl_{s_j}}^{(b_j)}(1-\eta_{jl_{s_j}}^{(h_j+1)})\Big)\times$$

$$\times\overline{g}_{I_{\bar d}J_{1,2}}(Y_{jL_{a_{h_j}}}^{(h_j+1)}, j\in J_1, \overline{Y}_{jL_{a_{h_j}}}^{(h_j+1)}, j\in J_2, \overline{Y}_{L_{\bar a_{h-1}}\setminus L_{\bar a_h}},\dots,\widetilde{X}_{I_{\bar d}\setminus R_{\bar k}}),$$

где J_1, J_2 – разбиение $\{1,\dots,c\}$, такое что J_1 содержит все $j\in\{1,\dots,c\}$, для которых берется $a_{h_j}0$, а J_2 – все j, для которых берется $a_{h_j}a_{h_j}$.

Тогда

$$I_{133}\stackrel{d}{=} it\sum_{D_+\ge 2}\sum_{\bar k: D_+>K_+\ge 1} E\Big[\exp\Big(it\Big(\sum_{j=1}^{c}Z_{n_jk_j}^{(1)}+\sum_{j=1}^{c}\sum_{r_j=0}^{k_j}Z_{n_jr_j}^{(2)}+$$

$$+\sum_{D_+\ge 2}W_{\bar d0}\Big)\Big)\sum_{h_j=0}^{k_j-1}\sum_{1\le a_{h_j}<a_{h_j-1}<\dots<k_j}\sum_{\bar A}W_{\bar d\bar k\bar A}\Big]=$$

$$= it\sum_{D_+\ge 2}\sum_{\bar k: D_+>K_+\ge 1}\sum_{h_j=0}^{k_j-1}\sum_{1\le a_{h_j}<a_{h_j-1}<\dots<k_j}\sum_{\bar A}\upsilon_{\bar A}=\sum_{\bar A}\Upsilon_{\bar A},$$

где

$$\upsilon_{\bar A}=E\Big[\exp\Big(it\Big(\sum_{j=1}^{c}Z_{n_jk_j}^{(1)}+\sum_{j=1}^{c}\sum_{r_j=0}^{k_j}Z_{n_jr_j}^{(2)}+\sum_{D_+\ge 2}W_{\bar d0}\Big)\Big)W_{\bar d\bar k\bar A}\Big],$$

$$\Upsilon_{\bar A}=it\sum_{D_+\ge 2}\sum_{\bar k: D_+>K_+\ge 1}\sum_{h_j=0}^{k_j-1}\sum_{1\le a_{h_j}<a_{h_j-1}<\dots<k_j}\upsilon_{\bar A}.$$

Рассмотрим одно из 2^c слагаемых, т.к. все они оцениваются одинаково:

$$\Upsilon_{\bar A}=it\sum_{D_+\ge 2}\sum_{\bar k: D_+>K_+\ge 1}\sum_{h_j=0}^{k_j-1}\sum_{1\le a_{h_j}<a_{h_j-1}<\dots<k_j}\times$$

$$\Big(E\Big[\exp\Big(it(\sum_{j=1}^{c}Z_{n_jk_j}^{(1)}+\sum_{j=1}^{c}\sum_{r_j=0}^{k_j}Z_{n_jr_j}^{(2)}\Big)\Big)\Big(\exp\Big(it\sum_{D_+\ge 2}W_{\bar d0}\Big)-1\Big)\Big)W_{\bar d\bar k\bar A}\Big]$$

$$+E\Big[\exp\Big(it\Big(\sum_{j=1}^{c}Z^{(1)}_{n_jk_j}+\sum_{j=1}^{c}\sum_{r_j=0}^{k_j}Z^{(2)}_{n_jr_j}\Big)\Big)W_{\bar d\bar k\bar A}\Big]\Big)=\Upsilon_{\bar A1}+\Upsilon_{\bar A2}.$$

Заметим, что $W_{\bar d\bar k\bar A}$ не зависит от

$$\sum_{j\in J_1}Z^{(2)}_{n_jh_j+1}(\overline{Y}_j^{(h_j+1)})+\sum_{j\in J_2}Z^{(1)}_{n_jh_j+1}(Y_j^{(h_j+1)})$$

при фиксированных α, η. Тогда имеем оценку

$$|\Upsilon_{\bar A1}|\le\bar\sigma_{\bar n}^{-1}|t|\sum_{D_+\ge2}\cdots\sum_{\mathcal{R}^{\bar k}_{\bar a_1}}E|\prod_{j=1}^{c}\prod_{s_j=1}^{a_{1j}}\alpha_{jl_{s_j}}\prod_{s_j=a_{1j}+1}^{k_j}\alpha_{jr_{s_j}}\prod_{s_j=k_j+1}^{d_j}(1-\alpha_{ji_{s_j}})$$

$$\times\prod_{s_j=a_{1j}+1}^{k_j}(1-\eta^{(1)}_{jr_{s_j}})\times\cdots\times\prod_{s_j=a_{hj-1}+1}^{a_{hj}}(\prod_{b_j=1}^{h_j-1}\eta^{(b_j)}_{jl_{s_j}}(1-\eta^{(h_j)}_{jl_{s_j}}))\times$$

$$\times\prod_{j\in J_1}(\prod_{s_j=1}^{a_{h_j}}\prod_{b_j=1}^{h_j+1}\eta^{(b_j)}_{jl_{s_j}})\prod_{j\in J_2}(\prod_{s_j=1}^{a_{h_j}}\prod_{b_j=1}^{h_j}\eta^{(b_j)}_{jl_{s_j}}(1-\eta^{(h_j+1)}_{jl_{s_j}}))\times\mathcal{A}\times\mathcal{B}\times\mathcal{C}|,$$

с условными математическими ожиданиями

$$\mathcal{A}=E[\exp(it(\sum_{j\in J_1}Z^{(2)}_{n_jh_j+1}+\sum_{j\in J_2}Z^{(1)}_{n_jh_j+1}))|\alpha,\eta],\tag{3}$$

$$\mathcal{B}=E[\exp(it(\sum_{j\in J_2}\sum_{s_j\ne\Theta_j}\prod_{b_j=1}^{h_j}\eta^{(b_j)}_{jl_{s_j}}(1-\eta^{(h_j+1)}_{jl_{s_j}})\alpha_{js_j}\overline{z}_{js_j}+$$

$$+\sum_{j\in J_1}\sum_{s_j\ne\Theta_j}\prod_{b_j=1}^{h_j+1}\eta^{(b_j)}_{jl_{s_j}}\alpha_{js_j}z_{js_j}))|\alpha,\eta],\tag{4}$$

$$\mathcal{C}=E[\prod_{j\in J_2}\prod_{s_j=\Theta_j}(\exp(it\prod_{b_j=1}^{h_j}\eta^{(b_j)}_{jl_{s_j}}(1-\eta^{(h_j+1)}_{js_j})\alpha_{js_j}\overline{z}_{js_j})-1)\times$$

$$\times\prod_{j\in J_1}\prod_{s_j=\Theta_j}(\exp(it\prod_{b_j=1}^{h_j+1}\eta^{(b_j)}_{jl_{s_j}}\alpha_{js_j}z_{js_j})-1)\times$$

$$\times(\exp(it\sum_{D_+\ge2}W_{\bar d0})-1)\sum_{\mathcal{I}^{\bar d}_{\bar k}}\overline{g}_{I_{\bar d}J_{1,2}}|\alpha,\eta],$$

где $\Theta_j = \bigcup\limits_{h_j} L_{a_{h_j}}^{a_{h_j}-1} \bigcup R_{a_{1_j}}^{k_j}$.

Сначала оценим $\mathcal{C}$. Применим неравенство Гельдера с $1 \le p \le 2$, $q = p/(p-1)$ и элементарные неравенства $|e^{ix}-1| \le 2|x|^{p-1}$, $|e^{ix}-1| \le |x|$:

$$\mathcal{C} \le (\prod_{j\in J_2}\prod_{s_j=\Theta_j} E[|\exp(it\prod_{b_j=1}^{h_j}\eta_{jl_{s_j}}^{(b_j)}(1-\eta_{js_j}^{(h_j+1)})\alpha_{js_j}\overline{z}_{js_j})-1|^q|\alpha,\eta]\times$$

$$\times\prod_{j\in J_1}\prod_{s_j=\Theta_j} E[|\exp(it\prod_{b_j=1}^{h_j+1}\eta_{jl_{s_j}}^{(b_j)}\alpha_{js_j}z_{js_j})-1|^q|\alpha,\eta]\times$$

$$\times E[|\exp(it\sum_{D_+\ge 2} W_{\bar{d}0})-1|^q|\alpha,\eta])^{\frac{1}{q}}\times$$

$$\times(2^{(2-p)(D_+-K_+)}\sum_{\mathcal{I}_{\bar{k}}^{\bar{d}}} E|\overline{g}_{I_{\bar{d}}J_{1,2}}|^p)^{\frac{1}{p}} \le$$

$$\le A_{\bar{m}}\,\overline{\sigma}_{\bar{n}}^{-p+1}\,|t|^{p-1+K_+}(\prod_{j\in J_2}\prod_{s_j=\Theta_j}\prod_{b_j=1}^{h_j}\eta_{jl_{s_j}}^{(b_j)q}(1-\eta_{js_j}^{(h_j+1)})^q\alpha_{js_j}^q E|z_{js_j}|^q\times$$

$$\times\prod_{j\in J_1}\prod_{s_j=\Theta_j}\prod_{b_j=1}^{h_j+1}\eta_{jl_{s_j}}^{(b_j)q}\alpha_{js_j}^q E|z_{js_j}|^q)^{\frac{1}{q}}\times$$

$$\times(\sum_{D_+\ge 2}\sum_{\mathcal{I}_{\bar{d}}}\prod_{j=1}^{c}\prod_{s_j=1}^{d_j}(1-\alpha_{ji_{s_j}})^p E|\overline{g}_{I_{\bar{d}}}|^p)^{\frac{1}{q}}(\sum_{\mathcal{I}_{\bar{k}}^{\bar{d}}} E|\overline{g}_{I_{\bar{d}}}|^p)^{\frac{1}{p}}.$$

Поскольку $q \ge 2$, имеем $E|z_{js_j}|^q \le Ez_{js_j}^2(2D_+)^{q-2}(\frac{\sigma_{\bar{n}}}{\overline{\sigma}_{\bar{n}}})^{q-2}$. Тогда

$$|\Upsilon_{\bar{A}1}| \le$$

$$A_{\bar{m}}\,\overline{\sigma}_{\bar{n}}^{-p}\,(\frac{\sigma_{\bar{n}}}{\overline{\sigma}_{\bar{n}}})^{q-2}\,|t|^p\sum_{D_+\ge 2}\sum_{\mathcal{L}_{\bar{a}_h}}\cdots\sum_{\mathcal{R}_{\bar{a}_1}^{\bar{k}}}|t|^{K_+}f(t)^{\frac{K_+}{2}}(E|\mathcal{A}|^2)^{\frac{1}{2}}(E|\mathcal{B}|^2)^{\frac{1}{2}}$$

$$\times(\prod_{j=1}^{c}\prod_{s_j=1}^{d_j}Ez_{js_j}^2\sum_{D_+\ge 2}\sum_{\mathcal{I}_{\bar{d}}}E|\overline{g}_{I_{\bar{d}}}|^p)^{\frac{1}{q}}(\sum_{\mathcal{I}_{\bar{k}}^{\bar{d}}}E|\overline{g}_{I_{\bar{d}}}|^p)^{\frac{1}{p}}.$$

По лемме 2 получаем оценки для множителей с $\mathcal{A}$ и $\mathcal{B}$:

$$E|\mathcal{A}|^2 \le |t|^{-2^{M_+-K_+}},$$

$$E|\mathcal{B}|^2 \le |t|^{-2^{M_+-K_+}(1-3K_+2^{-2M_+/3-2})}.$$

Кроме того, по неравенству Гельдера для сумм

$$\sum_{\mathcal{I}_{\bar d}} \prod_{j=1}^{c} \prod_{s_j=i_{j_1}}^{i_{jd_j}} Ez_{js_j}^2 \le (\sum_{j=1}^{c} \sum_{i_j=1}^{n_j} Ez_{ji_j}^2)^{D_+} = 1.$$

В результате получаем

$$|\Upsilon_{\bar A1}| \le$$
$$A_{\bar m}\,\overline{\sigma}_{\bar n}^{-p}\,(\frac{\sigma_{\bar n}}{\overline{\sigma}_{\bar n}})^{q-2}\,|t|^{p-2^{M_+/3-1}(2^{2M_+/3}-M_+/4)}\,\log^{\frac{1}{2}}(t)\sum_{D_+\ge 2}\sum_{\mathcal{I}_{\bar k}^{\bar d}} E|\overline{g}_{I_{\bar d}}|^p,$$

при $1 \le p \le 2$.

Для $\Upsilon_{\bar A2}$ запишем, использовав свойство полной вырожденности канонических функций,

$$|\Upsilon_{\bar A2}| \le$$
$$|t|\,\overline{\sigma}_{\bar n}^{-1}\sum_{D_+\ge 2}\cdots\sum_{\mathcal{I}_{\bar k}^{\bar d}} E|\prod_{j=1}^{c}\prod_{s_j=1}^{a_{1_j}}\alpha_{jl_{s_j}}\cdots\prod_{j\in J_2}(\prod_{s_j=1}^{a_{h_j}}\prod_{b_j=1}^{h_j}\eta_{jl_{s_j}}^{(b_j)}(1-\eta_{jl_{s_j}}^{(h_j+1)}))$$
$$\times\mathcal{A}\times\mathcal{B}\times\mathcal{D}\times\mathcal{E}|,$$

где $\mathcal{A}$, $\mathcal{B}$ определены в (3) и (4), соответственно.

$$\mathcal{D} = E[\exp(it\sum_{j=1}^{c}\sum_{s_j\ne I_{k_j}^{d_j}}(1-\alpha_{js_j})z_{js_j})|\alpha,\eta],$$

$$\mathcal{E} = E[\prod_{j\in J_2}\prod_{s_j=\Theta_j}(\exp(it\prod_{b_j=1}^{h_j}\eta_{jl_{s_j}}^{(b_j)}(1-\eta_{jl_{s_j}}^{(h_j+1)})\alpha_{js_j}\widehat{\widehat{z}}_{js_j})-1)\times$$

$$\times\prod_{j\in J_1}\prod_{s_j=\Theta_j}(\exp(it\prod_{b_j=1}^{h_j+1}\eta_{jl_{s_j}}^{(b_j)}\alpha_{js_j}\widehat{z}_{js_j})-1)\times$$

$$\times\prod_{j=1}^{c}\prod_{s_j=I_{k_j}^{d_j}}(\exp(it(1-\alpha_{js_j})\widehat{\widetilde{z}}_{js_j})-1)\overline{g}_{I_{\bar d}J_{1,2}}|\alpha,\eta].$$

Тогда по аналогии с предыдущими рассуждениями и по лемме 2

$$|\Upsilon_{\bar A2}|\le|t|^{2^{M_+}}\exp\Big(-\frac{t^2}{2}(1/3-(M_+-1)2^{-2M_+/3-2})\Big)\log(t)\times$$

$$\times\overline{\sigma}_{\bar n}^{-1}\sum_{D_+\ge2}\sum_{\bar k:D_+>K_+\ge1}\sum_{\mathcal{I}_{\bar d}}E|\prod_{j=1}^{c}\prod_{s_j=1}^{d_j}\widehat{z}_{j\,i_{s_j}}\overline{g}_{I_{\bar d}}|.$$

Остается оценить I_2. Запишем разложение

$$I_2\stackrel{d}{=}E\exp(it(Z_{\bar n}^{(1)}+Z_{\bar n}^{(2)}+\sum_{D_+\ge2}\sum_{\bar k=\bar0}^{\bar d}W_{\bar d\bar k}+\sum_{B_+\ge2}\sum_{D_+>B_+}\sum_{\bar k=\bar0}^{\bar b}V_{\bar d\bar b\bar k}))$$

$$-E\exp(it(Z_{\bar n}^{(1)}+Z_{\bar n}^{(2)}+\sum_{D_+\ge2}\sum_{\bar k=\bar0}^{\bar d}W_{\bar d\bar k}))=$$

$$=E[\exp(it(Z_{\bar n}^{(1)}+Z_{\bar n}^{(2)}+\sum_{D_+\ge2}\sum_{\bar k=\bar0}^{\bar d}W_{\bar d\bar k}+\sum_{B_+\ge2}\sum_{D_+>B_+}\sum_{K_+\ge1}V_{\bar d\bar b\bar k}))\times$$

$$\times(\exp(it\sum_{B_+\ge2}\sum_{D_+>B_+}V_{\bar d\bar b0})-1)]+$$

$$+E[\exp(it(Z_{\bar n}^{(1)}+Z_{\bar n}^{(2)}+\sum_{D_+\ge2}\sum_{\bar k=\bar0}^{\bar d}W_{\bar d\bar k}))\times$$

$$\times(\exp(it\sum_{B_+\ge2}\sum_{D_+>B_+}\sum_{K_+\ge1}V_{\bar d\bar b\bar k})-1)]=I_{21}+I_{22}.$$

По неравенству Гельдера

$$|I_{21}|\le$$

$$(E|E[\exp(itZ_{\bar n}^{(1)}|\alpha]|^2)^{\frac12}(E|E[\exp(it\sum_{B_+\ge2}\sum_{D_+>B_+}V_{\bar d\bar b0})-1|\alpha]|^2)^{\frac12}\le$$

$$\le|t|^{1-2^{M_+-2}}\sum_{B_+\ge2}\sum_{D_+>B_+}\sum_{\mathcal{L}_{\bar b}}E|\overline{g}_{I_{\bar d}L_{\bar b}}|.$$

Для I_{22} запишем

$$|I_{22}| \le E|\exp(it\sum_{B_+\ge 2}\sum_{D_+>B_+}\sum_{K_+\ge 1} V_{\bar d\bar b\bar k}) - 1| \le$$

$$\le \overline{\sigma}_{\bar n}^{-1}|t|\sum_{B_+\ge 2}\sum_{D_+>B_+}\sum_{K_+\ge 1}\prod_{j=1}^{c}\binom{b_j}{k_j} f(t)^{K_+}\sum_{\mathcal{L}_{\bar b}} E|\overline{g}_{I_{\bar d}L_{\bar b}}| \le$$

$$\le 3\cdot 2^{2M_+}|t|^{-1}\log(t)\overline{\sigma}_{\bar n}^{-1}\sum_{B_+\ge 2}\sum_{D_+>B_+}\sum_{\mathcal{L}_{\bar b}} E|\overline{g}_{I_{\bar d}L_{\bar b}}|.$$

Лемма 3. *Верны следующие соотношения:*

$$\left|\frac{\overline{\sigma}_{\bar n}^2}{\sigma_{\bar n}^2} - 1\right| \le 2\sigma_{\bar n}^{-2}\sum_{j=1}^{c}\sum_{l_j=1}^{n_j} Eg_{l_j}^2 I(|g_{l_j}| > \sigma_{\bar n}) +$$

$$+2M_{\mathsf{X}}(M_{\mathsf{X}}+4)\sigma_{\bar n}^{-1}\sum_{D_+\ge 2}\sum_{\mathcal{I}_{\bar d}} E|g_{I_{\bar d}}|I(|g_{I_{\bar d}}| > \sigma_{\bar n}),$$

$$\overline{\sigma}_{\bar n}^2 \ge 0.5\sigma_{\bar n}^2.$$

$$\sigma_{\bar n}^{-1}\sum_{D_+\ge 2}\sum_{\mathcal{I}_{\bar d}} E|\prod_{j=1}^{c}\prod_{k_j=1}^{d_j}\hat z_{j\,i_{k_j}}\overline{g}_{I_{\bar d}}| \le$$

$$\le A_{\bar m}(2^{m/2}\sum_{D_+\ge 2}\sigma_{\bar n}^{-1-D_+}\sum_{\mathcal{I}_{\bar d}} E|\prod_{j=1}^{c}\prod_{s_j=1}^{k_j}\widetilde{g}_{i_{s_j}}\widetilde{g}_{I_{\bar d}}| +$$

$$+\sigma_{\bar n}^{-2}\sum_{D_+\ge 2}\sum_{\mathcal{I}_{\bar d}} Eg_{I_{\bar d}}^2 I(|g_{I_{\bar d}}| \le \sigma_{\bar n}) +$$

$$+\sum_{D_+\ge 1}(\sigma_{\bar n}^{-1}\sum_{R_+\ge 2}\sum_{\mathcal{I}_{\bar r}} E|g_{I_{\bar r}}|I(|g_{I_{\bar r}}| > \sigma_{\bar n}))^{D_+}).$$

Доказательство. Является распространением лемм 3 и 4 для одновыборочных U-статистик из [3] на случай многовыборочных.

Объединяя формулы (1)–(2), оценки ε_1, ε_2, ε_3, лемму 3 и, используя соотношения

$$\sigma_{\bar n}^{-2}Eg_{I_{\bar d}}^2 I(|g_{I_{\bar d}}| \le \sigma_{\bar n}) = \sigma_{\bar n}^{-p}E|g_{I_{\bar d}}|^p I(|g_{I_{\bar d}}| \le \sigma_{\bar n}), \quad 1\le p\le 2,$$

$$\sigma_{\bar n}^{-1} E|g_{i_j}| I(|g_{i_j}| > \sigma_{\bar n}) = \sigma_{\bar n}^{-2} E g_{i_j}^2 I(|g_{i_j}| > \sigma_{\bar n}),$$

получим теорему 1.

Доказательство следствия 1. Для $\epsilon > 0$ можно записать

$$\sigma_{\bar n}^{-3} \sum_{j=1}^{c} \sum_{i_j=1}^{n_j} E|g_{i_j}|^3 I(|g_{i_j}| \le \sigma_{\bar n}) =$$

$$= \sigma_{\bar n}^{-3} \sum_{j=1}^{c} \sum_{i_j=1}^{n_j} E|g_{i_j}|^3 I(\epsilon\sigma_{\bar n} < |g_{i_j}| \le \sigma_n) +$$

$$+ \sigma_{\bar n}^{-3} \sum_{j=1}^{c} \sum_{i_j=1}^{n_j} E|g_{i_j}|^3 I(|g_{i_j}| \le \epsilon\sigma_{\bar n}) \le$$

$$\le \Lambda_{1,\bar n}(\epsilon) + \epsilon\sigma_{\bar n}^{-2} \sum_{j=1}^{c} \sum_{i_j=1}^{n_j} E g_{i_j}^2 = \Lambda_{1,n}(\epsilon) + \epsilon.$$

Далее,

$$\sigma_{\bar n}^{-1} \sum_{\mathcal{I}_{\bar d}} E|g_{I_{\bar d}}| I(|g_{I_{\bar d}}| > \sigma_{\bar n}) \le$$

$$\le \sigma_{\bar n}^{-\gamma_{D+}} \sum_{\mathcal{I}_{\bar d}} E|g_{I_{\bar d}}|^{\gamma_{D+}} I(|g_{I_{\bar d}}| > \sigma_{\bar n}) = \Lambda_{D_+,\bar n}(1).$$

$$\sum_{D_+ \ge 2} \sigma_{\bar n}^{-D_+ - 1} \sum_{\mathcal{I}_{\bar d}} E| \prod_{j=1}^{c} \prod_{s_j=1}^{d_j} \widetilde{g}_{i_{s_j}} \widetilde{g}_{I_{\bar d}}| \le$$

$$\le \sum_{D_+ \ge 2} (\sigma_{\bar n}^{-2} \sum_{\mathcal{I}_{\bar d}} E g_{I_{\bar d}}^2 I(|g_{I_{\bar d}}| \le \sigma_{\bar n}))^{1/2}.$$

И при $4/3 < p \le 2$ получаем

$$\sigma_{\bar n}^{-p} \sum_{\mathcal{I}_{\bar d}} E|g_{I_{\bar d}}|^p I(|g_{I_{\bar d}}| \le \sigma_{\bar n}) \le \Lambda_{D_+,\bar n}(\epsilon) + \Lambda_{D_+,\bar n}(0)\epsilon^{p-\gamma_{D+}}.$$

Таким образом, в теореме 1 для всякого $\epsilon > 0$

$$\Delta_{\bar n} \le C(\epsilon + \Lambda_{D_+,\bar n}(0)\epsilon^{p-\gamma_{D+}} + \epsilon^{1/3} + \Lambda_{1,\bar n}(\epsilon) +$$

$$+ \sum_{D_+ \ge 2} (\Lambda_{D_+,\bar n}(1) + \Lambda_{D_+,\bar n}(\epsilon) + \sqrt{\Lambda_{D_+,\bar n}(\epsilon)})) \to 0$$

при $\min\{n_1,\dots,n_c\}\to\infty$.

Доказательство следствия 2. Поскольку

$$\prod_{j=1}^{c}\prod_{s_j=1}^{d_j}|\widetilde{g}_{i_{s_j}}|^{1-2/D_+}\le\sigma_{\bar n}^{D_+-2},$$

имеем

$$\sigma_{\bar n}^{-D_+-1}\sum_{\mathcal{I}_{\bar d}}E|\prod_{j=1}^{c}\prod_{s_j=1}^{d_j}\widetilde{g}_{i_{s_j}}\widetilde{g}_{I_{\bar d}}|\le\sigma_{\bar n}^{-3}\sum_{\mathcal{I}_{\bar d}}E(\prod_{j=1}^{c}\prod_{s_j=1}^{d_j}|\widetilde{g}_{i_{s_j}}|^{2/D_+}\cdot|\widetilde{g}_{I_{\bar d}}|).$$

Используя неравенство $a\cdot b\le p^{-1}a^p+q^{-1}b^q$, $a>0$, $b>0$, $p^{-1}+q^{-1}=1$, с $p=\beta_{D_+}$ и $q=\frac{2D_++1}{2}$, получаем

$$\sigma_{\bar n}^{-3}\sum_{D_+\ge2}\sum_{\mathcal{I}_{\bar d}}E(\prod_{j=1}^{c}\prod_{s_j=1}^{d_j}|\widetilde{g}_{i_{s_j}}|^{2/D_+}\cdot|\widetilde{g}_{I_{\bar d}}|)\le$$

$$\le\sum_{D_+\ge2}\frac{1}{\beta_{D_+}}\sigma_{\bar n}^{-\beta_{D_+}}\sum_{\mathcal{I}_{\bar d}}E|\widetilde{g}_{I_{\bar d}}|^{\beta_{D_+}}+$$

$$+\sum_{D_+\ge2}\frac{2}{2D_++1}\sigma_{\bar n}^{-2D_+-1}\sum_{\mathcal{I}_{\bar d}}\prod_{j=1}^{c}\prod_{s_j=1}^{d_j}E|\widetilde{g}_{i_{s_j}}|^{\frac{2D_++1}{D_+}}.$$

Далее второе слагаемое оценивается следующим образом

$$\sum_{D_+\ge2}\frac{2}{2D_++1}(\sigma_{\bar n}^{-\frac{2D_++1}{D_+}}\sum_{j=1}^{c}\sum_{i_j=1}^{n_j}E\{|\widetilde{g}_{i_j}|^{3/D_+}|\widetilde{g}_{i_j}|^{2(D_+-1)/D_+}\})^{D_+}\le$$

$$\le\sum_{D_+\ge2}\frac{2}{2D_++1}(\sigma_{\bar n}^{-\frac{2D_++1}{D_+}}(\sum_{j=1}^{c}\sum_{i_j=1}^{n_j}E\widetilde{g}_{i_j}^2)^{\frac{D_+-1}{D_+}}(\sum_{j=1}^{c}\sum_{i_j=1}^{n_j}E|\widetilde{g}_{i_j}|^3)^{\frac{1}{D_+}})^{D_+}\le$$

$$\le A_{\bar m}\,\sigma_{\bar n}^{-3}\sum_{j=1}^{c}\sum_{i_j=1}^{n_j}E|\widetilde{g}_{i_j}|^3,$$

что дает утверждение следствия.

Литература

1. Ю. В. Боровских, *О нормальной аппроксимации U-статистик.* — Теор. вер. и ее прим. **45**, No. 3 (2000), 469–488.
2. Ю. В. Боровских, В. С. Королюк, *Теория U-статистик.* — Наукова думка, Киев (1994).
3. Л. В. Гадасина, *Граница Берри–Эссеена для U-статистик.* — Зап. научн. семин. ПОМИ. **298** (2003), 54–78.
4. D. Fraser, *Nonparametric methods in statistics.* — New York, Wiley (1957).
5. W. Hoeffding, *A class of statistics with asymptotically normal distribution.* — Ann. Math. Statist. **19**, No. 3 (1948), 293–325.
6. P. Sen, Sequential Nonparametrics: invariance principles and statistical inference. — New York, John Wiley and Sons (1981).
7. В. Феллер, *Введение в теорию вероятностей и ее приложения.* — Т. II, пер. с англ., Мир, М. (1971).

Gadasina L.V. A Berry–Esseen type estimations for multi-sample U-statistics.

The rate of convergence in the central limit theorem for nondegenerate multi-sample U-statistics of series of independent samples of independent random variables is investigated under minimal sufficient moment conditions on canonical functions of Hoeffding representation.

С.-Петербургский Университет
Путей Сообщения

Поступило 11 ноября 2005 г.