Пусть $\ell_1 = \frac{\ell\ell}{L},\ k_1 = \frac{\ell k}{L},\ \ell_1, k_1 \in \mathbb{Z}$, то есть обучающая выборка X длиной ℓ нацело разбивается в той же пропорции, что и генеральная выборка $\mathbb{X}$ длиной $L = \ell + k$.

Всюду далее верхние индексы выборок будут указывать на их мощность.

Зафиксируем конкретное разбиение генеральной выборки $\mathbb{X} = X \cup \bar{X}$ и докажем следующее утверждение:

Утверждение 0.1.

$$\frac{n(a,\bar{X})}{k} - \frac{n(a,X)}{\ell} = \frac{1}{C_k^{k_1} C_\ell^{\ell_1}} \sum_{\substack{X^{k_1}\subset\bar{X},\\ X^{\ell_1}\subset X}} \left(\frac{n(a,X^{k_1})}{k_1} - \frac{n(a,X^{\ell_1})}{\ell_1} \right).$$

Здесь сумма берётся по всевозможным выборкам $X^{k_1} \subset \bar{X}$ и $X^{\ell_1} \subset X$.

Доказательство.

$$\frac{1}{C_k^{k_1} C_\ell^{\ell_1}} \sum_{\substack{X^{k_1}\subset\bar{X},\\ X^{\ell_1}\subset X}} \left(\frac{n(a,X^{k_1})}{k_1} - \frac{n(a,X^{\ell_1})}{\ell_1} \right) = \frac{1}{C_k^{k_1} C_\ell^{\ell_1}} \left(C_{k-1}^{k_1-1} C_\ell^{\ell_1} \frac{n(a,\bar{X})}{k_1} - C_{\ell-1}^{\ell_1-1} C_k^{k_1} \frac{n(a,X)}{\ell_1} \right) =$$

$$= \frac{C_{k-1}^{k_1-1}}{C_k^{k_1}} \left(\frac{n(a,\bar{X})}{k_1} - \frac{C_{\ell-1}^{\ell_1-1} C_k^{k_1}}{C_{k-1}^{k_1-1} C_\ell^{\ell_1}} \frac{n(a,X)}{\ell_1} \right) = \frac{L}{\ell} \frac{C_{k-1}^{k_1-1}}{C_k^{k_1}} \left(\frac{n(a,\bar{X})}{k} - \frac{C_{\ell-1}^{\ell_1-1} C_k^{k_1}}{C_{k-1}^{k_1-1} C_\ell^{\ell_1}} \frac{n(a,X)}{\ell} \right).$$

Остается заметить, что

$$\frac{L}{\ell} \frac{C_{k-1}^{k_1-1}}{C_k^{k_1}} = 1 \quad \text{и} \quad \frac{C_{\ell-1}^{\ell_1-1} C_k^{k_1}}{C_{k-1}^{k_1-1} C_\ell^{\ell_1}} = 1.$$

■

Воспользовавшись этим утверждением, докажем следующую лемму, связывающую математическое ожидание большого равномерного уклонения с Радемахеровским средним (Rademacher averages) класса A.

Лемма 0.1.

$$\mathrm{Q}_{l,k}(A) = \frac{1}{C_L^\ell} \sum_{\mathbb{X}=X\cup\bar{X}} \sup_{a\in A} \left(\frac{n(a,\bar{X})}{k} - \frac{n(a,X)}{\ell} \right) \leqslant$$

$$\leqslant \frac{1}{C_L^\ell} \sum_{\mathbb{X}=X\cup\bar{X}} \frac{1}{C_\ell^{\ell_1}} \sum_{X=X^{\ell_1}\cup X^{k_1}} \sup_{a\in A} \left(\frac{n(a,X^{k_1})}{k_1} - \frac{n(a,X^{\ell_1})}{\ell_1} \right) = \mathrm{R}_{\ell,k}^\ell(A).$$

Доказательство.

$$\frac{1}{C_L^\ell}\sum_{\mathbb{X}=X\cup\bar{X}}\sup_{a\in A}\left(\frac{n(a,\bar{X})}{k}-\frac{n(a,X)}{\ell}\right)=$$

Утверждение 1

$$=\frac{1}{C_L^\ell}\sum_{\mathbb{X}=X\cup\bar{X}}\sup_{a\in A}\frac{1}{C_k^{k_1}C_\ell^{\ell_1}}\sum_{\substack{X^{k_1}\subset\bar{X},\\ X^{\ell_1}\subset X}}\left(\frac{n(a,X^{k_1})}{k_1}-\frac{n(a,X^{\ell_1})}{\ell_1}\right)\leqslant$$

выносим знак суммы из-под супремума

$$\leqslant\frac{1}{C_L^\ell}\frac{1}{C_k^{k_1}C_\ell^{\ell_1}}\sum_{\mathbb{X}=X\cup\bar{X}}\sum_{\substack{X^{k_1}\subset\bar{X},\\ X^{\ell_1}\subset X}}\left(\frac{n(a,X^{k_1})}{k_1}-\frac{n(a,X^{\ell_1})}{\ell_1}\right)=$$

каждую пару (X^{k_1},X^{ℓ_1}) мы учтём ровно $C_k^{k_1}$ раз

$$=\frac{1}{C_L^\ell}\frac{1}{C_k^{k_1}C_\ell^{\ell_1}}C_k^{k_1}\sum_{X^{k_1}\cap X^{\ell_1}=\emptyset}\left(\frac{n(a,X^{k_1})}{k_1}-\frac{n(a,X^{\ell_1})}{\ell_1}\right)=$$

$$=\frac{1}{C_L^\ell}\sum_{\mathbb{X}=X\cup\bar{X}}\frac{1}{C_\ell^{\ell_1}}\sum_{X=X^{\ell_1}\cup X^{k_1}}\left(\frac{n(a,X^{k_1})}{k_1}-\frac{n(a,X^{\ell_1})}{\ell_1}\right)$$

■