Overlapping Criteria with Multinomial Model

Bowen Deng

$$I_i \sim \frac{p^{x_{k+1}+\cdots+x_n}}{A(n,k,l)}, i = 1, 2, \cdots, t. \tag{1}$$

$$x_i \in \{0, 1\}. x_1 + x_2 + \cdots + x_n = l. \tag{2}$$

$$A(n,k,l) = \sum_{x_1+\cdots+x_n=l} p^{x_{k+1}+\cdots+x_n}. \tag{3}$$

$$p < k/2n. \tag{4}$$

$$
\begin{aligned}
l\mathbb{E}C(l) &= \mathbb{E}|I_1 \cap J_1| \\
&= \sum_{I_1} \Pr(I_1)\mathbb{E}(|I_1 \cap J_1| \big| I_1) \\
&= \sum_{u=0}^{l\wedge k} p^{l-u} C_k^u C_{n-k}^{l-u} / A(n,k,l)\mathbb{E}(|I_1 \cap J_1|)(I_1 = \{1,2,\cdots,u,k+1,\cdots,k+l-u\}) \\
&= \sum_{u=0}^{l\wedge k} p^{l-u} C_k^u C_{n-k}^{l-u} / A(n,k,l)\mathbb{E} \sum_{v=1,2,\cdots,u,k+1,\cdots,k+l-u} I(v \in J_1) \\
&= \sum_{u=0}^{l\wedge k} p^{l-u} C_k^u C_{n-k}^{l-u} / A(n,k,l) \sum_{v=1,2,\cdots,u,k+1,\cdots,k+l-u} \mathbb{E}I(v \in J_1) \\
&= \sum_{u=0}^{l\wedge k} p^{l-u} C_k^u C_{n-k}^{l-u} / A(n,k,l) \sum_{v=1,2,\cdots,u,k+1,\cdots,k+l-u} \Pr(v \in J_1) \\
&= \sum_{u=0}^{l\wedge k} p^{l-u} C_k^u C_{n-k}^{l-u} / A(n,k,l)(u\Pr(1 \in J_1) + (l-u)\Pr(k+1 \in J_1)) \\
&= \sum_{u=0}^{l\wedge k} p^{l-u} C_k^u C_{n-k}^{l-u} / A(n,k,l)(u\frac{A(n-1,k-1,l-1)}{A(n,k,l)} + (l-u)\frac{pA(n-1,k,l-1)}{A(n,k,l)}) \\
&= (\sum_{u=0}^{l\wedge k} up^{l-u} C_k^u C_{n-k}^{l-u})\frac{A(n-1,k-1,l-1)}{A^2(n,k,l)} + (\sum_{u=0}^{l\wedge k}(l-u)p^{l-u+1} C_k^u C_{n-k}^{l-u})\frac{A(n-1,k,l-1)}{A^2(n,k,l)} \\
&\triangleq s_1\frac{A(n-1,k-1,l-1)}{A^2(n,k,l)} + s_2\frac{A(n-1,k,l-1)}{A^2(n,k,l)} \\
s_1 &= k\sum_{u=1}^{l} p^{l-u} C_{k-1}^{u-1} C_{n-k}^{l-u} \\
&= k(\sum_{u=1}^{l} C_{k-1}^{u-1} p^{l-u} C_{n-k}^{l-u}) \\
&= kA(n-1,k-1,l-1) \\
s_2 &= (n-k)p^2 A(n-1,k,l-1) \\
l\mathbb{E}C(l) &= k(\frac{A(n-1,k-1,l-1)}{A(n,k,l)})^2 + (n-k)(p\frac{A(n-1,k,l-1)}{A(n,k,l)})^2
\end{aligned}
$$

Note that

$$
\begin{aligned}
\frac{\mathbb{E}(x_{k+1}+\cdots+x_n)}{\mathbb{E}(x_1+\cdots+x_k)} &= \frac{(n-k)\mathbb{E}x_{k+1}}{k\mathbb{E}x_1} \\
&= \frac{(n-k)\Pr(x_{k+1}=1)}{k\Pr(x_1=1)} \\
&= \frac{(n-k)\Pr(k+1\in I)}{k\Pr(1\in I)} \\
&= \frac{(n-k)pA(n-1,k,l-1)/A(n,k,l)}{kA(n-1,k-1,l-1)/A(n,k,l)} \\
&= \frac{(n-k)pA(n-1,k,l-1)}{kA(n-1,k-1,l-1)} \\
&\geqslant \frac{(n-k)p}{k}
\end{aligned}
$$

We expect $\rho \triangleq \frac{\mathbb{E}(x_{k+1}+\cdots+x_n)}{\mathbb{E}(x_1+\cdots+x_k)}$ to be smaller than $\frac{1}{2}$. Therefore, $p < \frac{k}{2n}$ is a reasonable assumption.

Lemma 1. If $f_1(x)$ is unimodal at a, $f_2(x)$ is unimodal at b, then $f_1(x)+f_2(x)$ takes maximum point in $[a,b]$.

Lemma 2. The following identities of $A(n,k,l)$ holds:

$$
\begin{aligned}
A(n,k,l) &= A(n-1,k-1,l-1) + A(n-1,k-1,l) \\
&= pA(n-1,k,l-1) + A(n-1,k,l)
\end{aligned} \tag{5}
$$