Lower Bounds on the Probability of a Finite Union of Events

Jun Yang
(joint work with Fady Alajaji and Glen Takahara)

Department of Mathematics and Statistics, Queen's University, Kingston, Canada

ISIT 2014

Outline

Problem Formulation

- Consider a finite family of events $A_1, \ldots, A_N$ in a finite probability space $(\Omega, \mathscr{F}, P)$, where N is a fixed positive integer.
- We are interested in lower bounds of $P\left(\bigcup_{i=1}^{N} A_i\right)$ in terms of the individual event probabilities $P(A_i)$'s and the pairwise event probabilities $P(A_i \cap A_j)$'s. For example,

$$P\left(\bigcup_{i=1}^{N} A_i\right) \geq \max_i P(A_i). \tag{1}$$

$$P\left(\bigcup_{i=1}^{N} A_i\right) \geq \sum_i P(A_i) - \sum_{i<j} P(A_i \cap A_j). \tag{2}$$

Problem Formulation

- Assume a vector θ represents partial information of $P\left(\bigcup_{i=1}^{N} A_i\right)$. That is, each element of θ equals to a (linear) function of $P(A_i)$'s and $P(A_i \cap A_j)$'s. For example,

$$\theta = (P(A_1), P(A_2), \ldots, P(A_N)). \tag{3}$$

$$\theta = \left(\sum_i P(A_i), \sum_{i<j} P(A_i \cap A_j)\right). \tag{4}$$

- Then a lower bound of $P\left(\bigcup_{i=1}^{N} A_i\right)$ is a function of θ, $\ell(\theta)$, such that

$$P\left(\bigcup_{i=1}^{N} A_i\right) \geq \ell(\theta), \tag{5}$$

for any $\{A_i\}$ that satisfy the partial information represented by θ.

Problem Formulation

- For a given definition of θ, for example, $\theta = (P(A_1), \cdots, P(A_N))$, there are many lower bounds that are functions of only θ:

$$\begin{aligned} P\left(\bigcup_{i=1}^{N} A_i\right) &\geq \theta_1 = P(A_1), \\ P\left(\bigcup_{i=1}^{N} A_i\right) &\geq \frac{\sum_i \theta_i}{N} = \frac{\sum_i P(A_i)}{N}, \\ P\left(\bigcup_{i=1}^{N} A_i\right) &\geq \max_i \theta_i = \max_i P(A_i). \end{aligned} \tag{6}$$

- What is the optimal lower bound in the class of lower bounds that are functions of θ?

Problem Formulation

- Let Θ denote the set of all possible values of θ (for a given definition of θ) and $\mathscr{L}_\Theta$ the set of all lower bounds on $P\left(\bigcup_{i=1}^N A_i\right)$ that are functions of only θ.

Definition

We say that a lower bound $\ell^\star \in \mathscr{L}_\Theta$ is optimal in $\mathscr{L}_\Theta$ if $\ell^\star(\theta) \geq \ell(\theta)$ for all $\theta \in \Theta$ and $\ell \in \mathscr{L}_\Theta$.

- Is $\ell(\theta) = \max_i \theta_i = \max_i P(A_i)$ optimal in the class of lower bounds that are functions of $\theta = (P(A_1), \ldots, P(A_N))$?
- How to prove a lower bound is optimal?

Problem Formulation

Definition

We say that a lower bound $\ell \in \mathscr{L}_\Theta$ is achievable if for every $\theta \in \Theta$,

$$\inf_{A_1,\dots,A_N} P\left(\bigcup_{i=1}^{N} A_i\right) = \ell(\theta), \tag{7}$$

where the infimum ranges over all collections $\{A_1, \dots, A_N\}$, $A_i \in \mathscr{F}$, such that $\{A_1, \dots, A_N\}$ is represented by θ.

Lemma

A lower bound $\ell^\star \in \mathscr{L}_\Theta$ is optimal in $\mathscr{L}_\Theta$ if and only if it is achievable.

Problem Formulation

We can therefore prove optimality by proving achievability:

- Step 1: prove $\ell(\theta)$ is a lower bound.
- Step 2: prove for any value of $\theta \in \Theta$, one can construct $\{A_i^*\}$ such that $P\left(\bigcup_i A_i^*\right) = \ell(\theta)$.

For example,

- $P\left(\bigcup_{i=1}^N A_i\right) \geq \max_i P(A_i)$ is the optimal lower bound in the class of lower bounds that are functions of $\theta = (P(A_1), \ldots, P(A_N))$.
- $P\left(\bigcup_{i=1}^N A_i\right) \geq \sum_i P(A_i) - \sum_{i<j} P(A_i \cap A_j)$ is not optimal lower bound in the class of lower bounds that are functions of $\theta = \left(\sum_i P(A_i), \sum_{i<j} P(A_i \cap A_j)\right)$.

Dawson-Sankoff (DS) Bound, 1967

- For each outcome $x \in \mathscr{F}$, let the degree of x, denoted by $\deg(x)$, be the number of A_i's that contain x.
- Define $a_k := P\left(\{x \in \bigcup_i A_i, \deg(x) = k\}\right)$, then one can verify

$$\begin{aligned} P\left(\bigcup_i A_i\right) &= \sum_{k=1}^{N} a_k, \\ \sum_i P(A_i) &= \sum_{k=1}^{N} k a_k, \\ \sum_{i<j} P(A_i \cap A_j) &= \sum_{k=2}^{N} \frac{k(k-1)}{2} a_k. \end{aligned} \tag{8}$$

Dawson-Sankoff (DS) Bound, 1967

- The Dawson-Sankoff (DS) bound is the solution of the following linear programming (LP) problem:

$$\min_{\{a_k \geq 0\}} \sum_{k=1}^{N} a_k, \quad \text{s.t.} \quad \begin{aligned} \sum_{k=1}^{N} k a_k &= \sum_i P(A_i), \\ \sum_{k=1}^{N} \frac{k(k-1)}{2} a_k &= \sum_{i<j} P(A_i \cap A_j). \end{aligned} \tag{9}$$

- The DS Bound is optimal in the class of lower bounds that are functions of $\theta = \left(\sum_i P(A_i), \sum_{i<j} P(A_i \cap A_j)\right) =: (\theta_1, \theta_2)$,

$$P\left(\bigcup_{i=1}^{N} A_i\right) \geq \frac{\kappa\theta_1^2}{(2-\kappa)\theta_1 + 2\theta_2} + \frac{(1-\kappa)\theta_1^2}{(1-\kappa)\theta_1 + 2\theta_2}, \tag{10}$$

where $\kappa = \frac{2\theta_2}{\theta_1} - \lfloor \frac{2\theta_2}{\theta_1} \rfloor$ and $\lfloor x \rfloor$ denotes the largest integer less than or equal to x.

Kuai-Alajaji-Takahara (KAT) Bound, 2000

- Define $a_i(k) = P(\{x \in A_i, \deg(x) = k)$.Recall that $a_k := P(\{x \in \bigcup_i A_i, \deg(x) = k\})$, one can verify

$$\sum_{i=1}^{N} a_i(k) = ka_k, \quad \Rightarrow \quad P\left(\bigcup_i A_i\right) = \sum_k a_k = \sum_k \sum_i \frac{a_i(k)}{k},$$
$$P(A_i) = \sum_{k=1}^{N} a_i(k), \quad \sum_{j:j\neq i} P(A_i \cap A_j) = \sum_{k=2}^{N} (k-1)a_i(k). \tag{11}$$

- The KAT bound is the solution of the following LP problem:

$$\min_{\{a_i(k)\geq 0\}} \sum_{k=1}^{N}\sum_{i=1}^{N} \frac{a_i(k)}{k}, \quad \text{s.t.} \quad \sum_{k=1}^{N} a_i(k) = P(A_i), \quad i = 1, \dots, N,$$
$$\sum_{k=1}^{N} (k-1)a_i(k) = \sum_{j:j\neq i} P(A_i \cap A_j), \quad i = 1, \dots, N. \tag{12}$$

Kuai-Alajaji-Takahara (KAT) Bound, 2000

- Let $\alpha_i := P(A_i)$, $\quad \gamma_i := \sum_j P(A_i \cap A_j) = P(A_i) + \sum_{j:j \neq i} P(A_i \cap A_j)$.
- The KAT bound,

$$P\left(\bigcup_{i=1}^{N} A_i\right) \geq \sum_{i=1}^{N} \left\{ \left[\frac{1}{\lfloor \frac{\gamma_i}{\alpha_i} \rfloor} - \frac{\frac{\gamma_i}{\alpha_i} - \lfloor \frac{\gamma_i}{\alpha_i} \rfloor}{(1 + \lfloor \frac{\gamma_i}{\alpha_i} \rfloor)(\lfloor \frac{\gamma_i}{\alpha_i} \rfloor)} \right] \alpha_i \right\}, \qquad (13)$$

  where $\lfloor x \rfloor$ is the largest positive integer less than or equal to x,
- is not optimal for $\theta = \Big(P(A_1), \ldots, P(A_N), \sum_{j:j \neq 1} P(A_1 \cap A_j), \ldots, \sum_{j:j \neq N} P(A_N \cap A_j)\Big)$.

New Lower Bounds which are sharper than KAT Bound

- Recall that $a_i(k) = P(\{x \in A_i, \deg(x) = k)$, then we observe

$$a_i(N) = P(\{x \in A_i, \deg(x) = N)$$

However, $\deg(x) = N \Leftrightarrow x \in A_i$ for all i, therefore

$$a_1(N) = a_2(N) = \cdots = a_N(N).$$

- Furthermore, by the definitions of $a_k := P(\{x \in \bigcup_i A_i, \deg(x) = k\})$ and $a_i(k)$, we observe that $a_k \geq a_i(k)$ for all i and all k. Also, since $a_k = \frac{\sum_i a_i(k)}{k}$, one can write

$$\frac{\sum_i a_i(k)}{k} \geq a_i(k)$$

for all i and all k.

- Note that when $k = N$, $\frac{\sum_i a_i(k)}{k} \geq a_i(k)$ reduces to $a_1(N) = a_2(N) = \cdots = a_N(N)$.

New analytical Lower Bound

The new analytical lower bound is the solution of the LP problem:

$$\min_{\{a_i(k)\geq 0\}} \sum_{k=1}^{N}\sum_{i=1}^{N}\frac{a_i(k)}{k}, \quad \text{s.t.} \quad \sum_{k=1}^{N} a_i(k) = P(A_i), \quad i=1,\dots,N,$$
$$\sum_{k=1}^{N}(k-1)a_i(k) = \sum_{j:j\neq i} P(A_i\cap A_j), \quad i=1,\dots,N, \tag{14}$$
$$a_1(N) = a_2(N) = \cdots = a_N(N).$$

The new analytical lower bound is given by

$$P\left(\bigcup_{i=1}^{N} A_i\right) \geq \delta + \sum_{i=1}^{N} \left\{\left[\frac{1}{\chi(\frac{\gamma_i'}{\alpha_i'})} - \frac{\frac{\gamma_i'}{\alpha_i'} - \chi(\frac{\gamma_i'}{\alpha_i'})}{[1+\chi(\frac{\gamma_i'}{\alpha_i'})][\chi(\frac{\gamma_i'}{\alpha_i'})]}\right]\alpha_i'\right\}, \tag{15}$$

where $\delta := \{\max_i [\gamma_i - (N-1)\alpha_i]\}^+ \geq 0$, $\alpha_i' := \alpha_i - \delta$, $\gamma_i' := \gamma_i - N\delta$, and

$$\chi(x) := \begin{cases} n-1 & \text{if } x = n \text{ where } n \geq 2 \text{ is a integer} \\ \lfloor x \rfloor & \text{otherwise} \end{cases}$$

New Analytical Lower Bound

- Let ℓ_{NEW} denote the new analytical bound and ℓ_{KAT} denote the KAT lower bound. Then the improvement of the new analytical bound over the existing KAT bound, i.e., $\ell_{\mathrm{NEW}} - \ell_{\mathrm{KAT}}$, satisfies the following inequality

$$\ell_{\mathrm{NEW}} - \ell_{\mathrm{KAT}} \geq \left\{ \sum_{i=1}^{N} \frac{\left[N - \chi(\frac{\gamma_i}{\alpha_i})\right]\left[N - \chi(\frac{\gamma_i}{\alpha_i}) - 1\right]}{\chi(\frac{\gamma_i}{\alpha_i})\left[\chi(\frac{\gamma_i}{\alpha_i}) + 1\right]} \right\} \frac{\delta}{N} \geq 0. \quad (16)$$

- The new analytical lower bound is still not optimal for $\theta = \Big(P(A_1), \ldots, P(A_N), \sum_{j:j\neq 1} P(A_1 \cap A_j), \ldots, \sum_{j:j\neq N} P(A_N \cap A_j)\Big)$.

New Optimal Lower Bound

- The solution of the following LP problem:

$$
\begin{aligned}
&\min_{\{a_i(k)\}} \sum_{k=1}^{N}\sum_{i=1}^{N} \frac{a_i(k)}{k}, \\
\text{s.t.} \quad & \sum_{k=1}^{N} a_i(k) = P(A_i), \quad i = 1, \dots, N, \\
& \sum_{k=1}^{N}(k-1)a_i(k) = \sum_{j:j\neq i} P(A_i \cap A_j), \quad i = 1, \dots, N, \\
& \frac{\sum_i a_i(k)}{k} \geq a_i(k), \quad i = 1, \dots, N, \quad k = 1, \dots, N, \\
& a_i(k) \geq 0, \quad k = 1, \dots, N, \quad i = 1, \dots, N.
\end{aligned} \tag{17}
$$

- is optimal in the class of lower bounds which are functions of $\theta = \Big(P(A_1), \dots, P(A_N), \sum_{j:j\neq 1} P(A_1 \cap A_j), \dots, \sum_{j:j\neq N} P(A_N \cap A_j)\Big)$.

New Optimal Lower Bound

- We gave a construction proof for the achievability in the paper.
- We could also obtain the optimal upper bound for $\theta = \left(P(A_1), \ldots, P(A_N), \sum_{j:j\neq 1} P(A_1 \cap A_j), \ldots, \sum_{j:j\neq N} P(A_N \cap A_j)\right)$.
- All bounds can be applied to any general probability of error estimation problem, including channel coding.

Numerical Examples

Table: Comparison of Lower Bounds.

System	$P\left(\bigcup_{i=1}^{N} A_i\right)$	DS	KAT	New Bound 1	New Bound 2
I	0.7890	0.7007	0.7247	0.7247	0.7487
II	0.6740	0.6150	0.6227	0.6227	0.6398
III	0.7890	0.6933	0.7222	0.7222	0.7427
IV	0.9687	0.8879	0.8909	0.8909	0.9044
V	0.3900	0.3800	0.3833	0.3900	0.3900
VI	0.3252	0.2706	0.2769	0.3205	0.3252
VII	0.5346	0.3989	0.4434	0.4562	0.5090
VIII	0.5854	0.5395	0.5412	0.5464	0.5513

References

- Dawson-Sankoff (DS) Bound, 1967
  D. Dawson, D. Sankoff, An inequality for probabilities, Proceedings of the American Mathematical Society 18 (3) (1967) 504-507.
- Kuai-Alajaji-Takahara (KAT) Bound, 2000
  H. Kuai, F. Alajaji, G. Takahara, A lower bound on the probability of a finite union of events, Discrete Mathematics 215 (1-3) (2000) 147-158.
- The New Bounds
  J. Yang, F. Alajaji, G. Takahara, Lower Bounds on the Probability of a Finite Union of Events, submitted. Online: http://arxiv.org/abs/1401.5543