Non-Shannon entropy inequalities and linear rank inequalities

Randall Dougherty
(Center for Communication Research, La Jolla)
Chris Freiling
(California State University, San Bernardino)
Ken Zeger
(University of California, San Diego)

Shannon inequalities

$$H(\emptyset) = 0$$

$$H(A) \geq 0$$

$$H(A|B) = H(AB) - H(B) \geq 0$$

$$I(A;B) = H(A) + H(B) - H(AB) \geq 0$$

$$I(A;B|C) = H(AC) + H(BC) - H(C) - H(ABC) \geq 0$$

A non-Shannon inequality (Zhang-Yeung)

$$I(A;B) \leq 2I(A;B|C) + I(A;C|B) + I(B;C|A) + I(A;B|D) + I(C;D)$$

Copying random variables

If A, B, and C are sets of random variables, then one can construct a new set of random variables R such that:

- (A, B) and (R, B) are identically distributed;
- $I(R; AC|B) = 0$.

We say that “R is a C-copy of A over B.”

$$p(A = a, B = b, C = c, R = r) = \frac{p(A = a, B = b, C = c)p(A = r, B = b)}{p(B = b)}$$

The Zhang-Yeung inequality

$$I(A;B) \leq 2I(A;B|C) + I(A;C|B) + I(B;C|A) + I(A;B|D) + I(C;D)$$

is a consequence of “R is a D-copy of C over AB”.

The Zhang-Yeung inequality

$$
\begin{aligned}
I(A;B) \le\ & 2I(A;B|C) + I(A;C|B) + I(B;C|A) + I(A;B|D) + I(C;D) \\
& + 3I(R;CD|AB) - (H(R) - H(C)) + 2(H(RA) - H(CA)) \\
& + 2(H(RB) - H(CB)) - 3(H(RAB) - H(CAB))
\end{aligned}
$$

is a Shannon inequality.

The Zhang-Yeung inequality

$$
\begin{aligned}
&I(C;D|R) + I(C;R|A) + I(C;R|B) + I(C;R|ABD) + I(D;R|A) + \\
&I(D;R|B) + I(D;R|ABC) + I(A;B|CR) + I(A;B|DR) + I(AB;R|CD) + \\
&\quad I(A;B) = 2I(A;B|C) + I(A;C|B) + I(B;C|A) + I(A;B|D) + I(C;D) \\
&\qquad + 3I(R;CD|AB) - (H(R) - H(C)) + 2(H(RA) - H(CA)) \\
&\qquad + 2(H(RB) - H(CB)) - 3(H(RAB) - H(CAB))
\end{aligned}
$$

is an entropy identity.

Procedure for generating inequalities

- List known inequalities
- Determine extreme rays of polytope
- Discard rays known to be entropic [skip this]
- List copy specifications
- Test extreme rays against copy specifications
- If a test fails, determine new inequality and start over

Inequalities obtained from two copy variables

$$
\begin{aligned}
2I(A;B) \le\ & 5I(A;B|C) + 3I(A;C|B) + I(B;C|A) \\
& + 2I(A;B|D) + 2I(C;D) \\
2I(A;B) \le\ & 4I(A;B|C) + 2I(A;C|B) + I(B;C|A) \\
& + 3I(A;B|D) + I(A;D|B) + 2I(C;D) \\
2I(A;B) \le\ & 4I(A;B|C) + 4I(A;C|B) + I(B;C|A) \\
& + 2I(A;B|D) + I(A;D|B) + I(B;D|A) + 2I(C;D) \\
2I(A;B) \le\ & 3I(A;B|C) + 3I(A;C|B) + 3I(B;C|A) \\
& + 2I(A;B|D) + 2I(C;D) \\
2I(A;B) \le\ & 3I(A;B|C) + 4I(A;C|B) + 2I(B;C|A) \\
& + 3I(A;B|D) + I(A;D|B) + 2I(C;D) \\
2I(A;B) \le\ & 3I(A;B|C) + 2I(A;C|B) + 2I(B;C|A) \\
& + 2I(A;B|D) + I(A;D|B) + I(B;D|A) + 2I(C;D)
\end{aligned}
$$

Common form of inequalities

$$
\begin{aligned}
aI(A;B) \le\ & bI(A;B|C) + cI(A;C|B) + dI(B;C|A) \\
& + eI(A;B|D) + fI(A;D|B) + gI(B;D|A) \\
& + hI(C;D) + iI(C;D|A)
\end{aligned}
$$

Inequalities obtained from three copy variables

$$
\begin{aligned}
2I(A;B) \le {} & 3I(A;B|C) + 3I(A;C|B) + 2I(B;C|A) \\
& + 2I(A;B|D) + 2I(C;D) \\
2I(A;B) \le {} & 5I(A;B|C) + 2I(A;C|B) + I(B;C|A) \\
& + 2I(A;B|D) + 2I(C;D) \\
2I(A;B) \le {} & 4I(A;B|C) + I(A;C|B) + 3I(B;C|A) \\
& + 2I(A;B|D) + 2I(C;D)
\end{aligned}
$$

etc.

Inequalities obtained from three copy variables

34 new inequalities (744 counting permutations)

plus 1 previous inequality (24 counting permutations)

Inequalities obtained from four copy variables

(using at most three copy steps)

203 new inequalities (4632 counting permutations)

plus 11 previous inequalities (264 counting permutations)

Generating new inequalities from known ones

$$2I(A;B) \leq 5I(A;B|C) + 3I(A;C|B) + I(B;C|A) + 2I(A;B|D) + 2I(C;D)$$

follows from "R is a D-copy of A over BC" and instances of the Zhang-Yeung inequality.

$$2I(A;B) \leq 5I(A;B|C) + 3I(A;C|B) + I(B;C|A) + 2I(A;B|D) + 2I(C;D)$$

follows from "R is a D-copy of A over BC" and the following instance of the Zhang-Yeung inequality:

$$\begin{aligned} I(AR;BR) \leq\ & 2I(AR;BR|CR) + I(AR;CR|BR) + I(BR;CR|AR) \\ & + I(AR;BR|D) + I(CR;D) \end{aligned}$$

$$\begin{aligned} 2I(A;B) \leq\ & 5I(A;B|C) + 3I(A;C|B) + I(B;C|A) + 2I(A;B|D) + 2I(C;D) \\ & + I(AR;BR) - 2I(AR;BR|CR) - I(AR;CR|BR) \\ & - I(BR;CR|AR) - I(AR;BR|D) - I(CR;D) \end{aligned}$$

follows from “R is a D-copy of A over BC”.

$$
\begin{aligned}
I(A;B) + I(A;B) \leq\ & 2I(A;B|C) + 2I(A;B|C) + I(A;B|C) \\
& + I(A;C|B) + I(A;C|B) + I(A;C|B) + I(B;C|A) \\
& + I(A;B|D) + I(A;B|D) + I(C;D) + I(C;D) \\
& + I(AR;BR) - 2I(AR;BR|CR) - I(AR;CR|BR) \\
& - I(BR;CR|AR) - I(AR;BR|D) - I(CR;D)
\end{aligned}
$$

follows from “R is a D-copy of A over BC”.

$$
\begin{aligned}
I(A;B) + I(A;B) \leq\ & 2I(A;B|C) + 2I(A;B|C) + I(A;B|C) \\
& + I(A;C|B) + I(A;C|B) + I(A;C|B) + I(B;C|A) \\
& + I(A;B|D) + I(A;B|D) + I(C;D) + I(C;D) \\
& + I(AR;BR) - 2I(AR;BR|CR) - I(AR;CR|BR) \\
& - I(BR;CR|AR) - I(AR;BR|D) - I(CR;D)
\end{aligned}
$$

follows from “R is a D-copy of A over BC”.

$$
\begin{aligned}
aI(A;B) + dI(A;B) \le\ & bI(A;B|C) + 2dI(A;B|C) + hI(A;B|C) \\
& + aI(A;C|B) + cI(A;C|B) + dI(A;C|B) + dI(B;C|A) \\
& + aI(A;B|D) + dI(A;B|D) + dI(C;D) + hI(C;D) \\
& + aI(AR;BR) - bI(AR;BR|CR) - cI(AR;CR|BR) \\
& - dI(BR;CR|AR) - aI(AR;BR|D) - hI(CR;D)
\end{aligned}
$$

follows from "R is a D-copy of A over BC".

$$
\begin{aligned}
(a+d)I(A;B) \le\ & (b+2d+h)I(A;B|C) + (a+c+d)I(A;C|B) + dI(B;C|A) \\
& + (a+d)I(A;B|D) + (d+h)I(C;D) \\
& + aI(AR;BR) - bI(AR;BR|CR) - cI(AR;CR|BR) \\
& - dI(BR;CR|AR) - aI(AR;BR|D) - hI(CR;D)
\end{aligned}
$$

follows from "R is a D-copy of A over BC".

$$(a+d)I(A;B) \leq (b+2d+h)I(A;B|C) + (a+c+d)I(A;C|B) + dI(B;C|A)$$
$$+ (a+d)I(A;B|D) + (d+h)I(C;D)$$

follows from "R is a D-copy of A over BC" and:

$$aI(AR;BR) \leq bI(AR;BR|CR) + cI(AR;CR|BR) + dI(BR;CR|AR)$$
$$+ aI(AR;BR|D) + hI(CR;D)$$

$$(a+d)I(A;B) \leq (b+2d+h)I(A;B|C) + (a+c+d)I(A;C|B) + dI(B;C|A)$$
$$+ (a+d)I(A;B|D) + (d+h)I(C;D)$$

follows from "R is a D-copy of A over BC" and an instance of:

$$aI(A;B) \leq bI(A;B|C) + cI(A;C|B) + dI(B;C|A)$$
$$+ aI(A;B|D) + hI(C;D)$$

If

$$aI(A;B) \leq bI(A;B|C) + cI(A;C|B) + dI(B;C|A) + aI(A;B|D) + hI(C;D)$$

is an information inequality, then

$$(a+d)I(A;B) \leq (b+2d+h)I(A;B|C) + (a+c+d)I(A;C|B) + dI(B;C|A) + (a+d)I(A;B|D) + (d+h)I(C;D)$$

is an information inequality.

If

$$\begin{aligned} aI(A;B) \le\ & bI(A;B|C) + cI(A;C|B) + dI(B;C|A) \\ & + aI(A;B|D) + fI(A;D|B) + gI(B;D|A) \\ & + hI(C;D) + iI(C;D|A) + jI(C;D|B) \end{aligned}$$

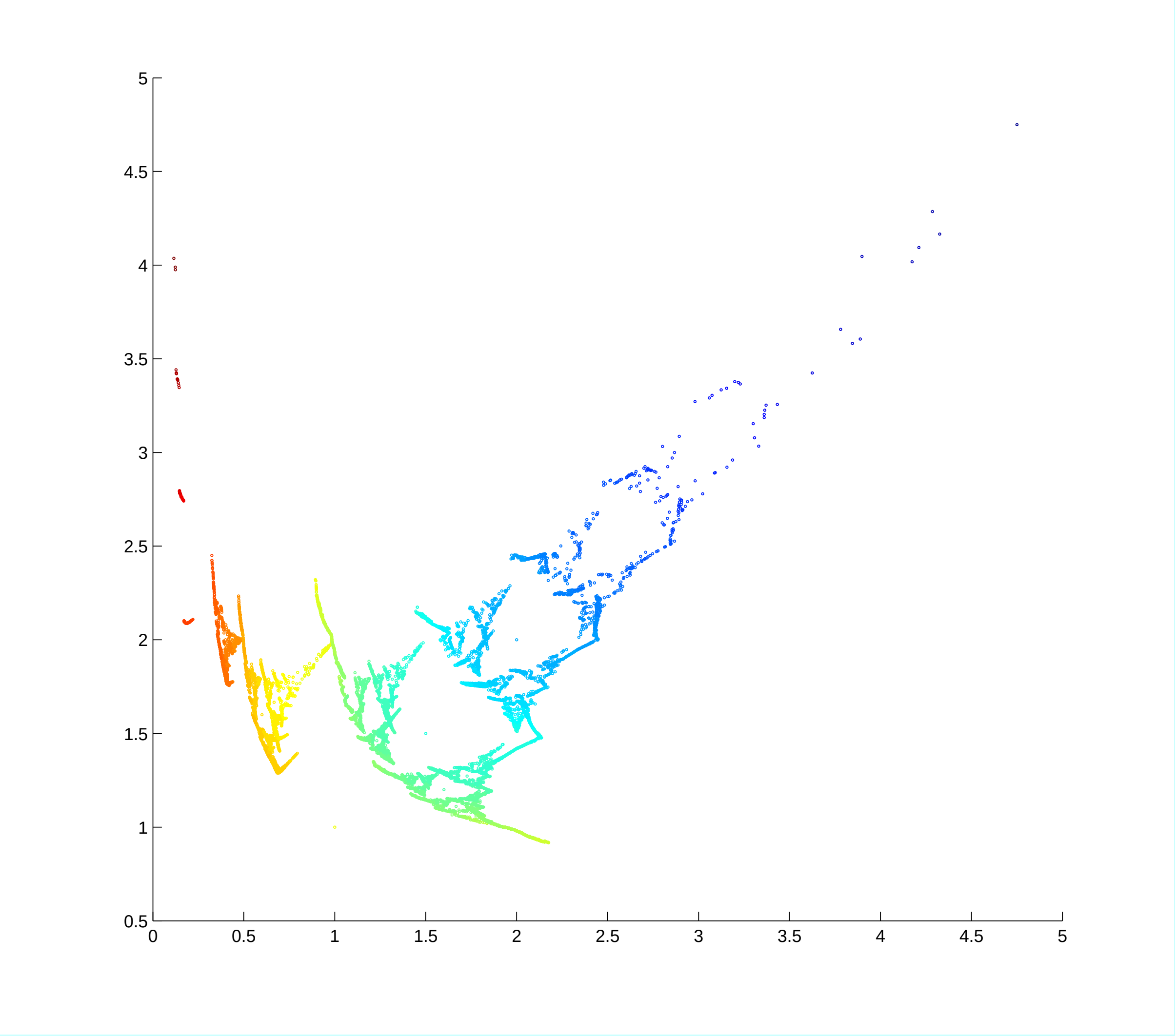
is an information inequality, then

$$\begin{aligned} (a+d)I(A;B) \le\ & (b+2d+h+i)I(A;B|C) + (a+c+d+f+g)I(A;C|B) \\ & + dI(B;C|A) + (a+d)I(A;B|D) + fI(A;D|B) + gI(B;D|A) \\ & + (d+h)I(C;D) + iI(C;D|A) + jI(C;D|B) \end{aligned}$$

is an information inequality.

Iterating this rule gives:

$$\begin{aligned}
I(A;B) &\leq 2I(A;B|C) + I(A;C|B) + I(B;C|A) + I(A;B|D) + I(C;D) \\
2I(A;B) &\leq 5I(A;B|C) + 3I(A;C|B) + I(B;C|A) + 2I(A;B|D) + 2I(C;D) \\
3I(A;B) &\leq 9I(A;B|C) + 6I(A;C|B) + I(B;C|A) + 3I(A;B|D) + 3I(C;D) \\
4I(A;B) &\leq 14I(A;B|C) + 10I(A;C|B) + I(B;C|A) + 4I(A;B|D) + 4I(C;D) \\
5I(A;B) &\leq 20I(A;B|C) + 15I(A;C|B) + I(B;C|A) + 5I(A;B|D) + 5I(C;D) \\
6I(A;B) &\leq 27I(A;B|C) + 21I(A;C|B) + I(B;C|A) + 6I(A;B|D) + 6I(C;D) \\
7I(A;B) &\leq 35I(A;B|C) + 28I(A;C|B) + I(B;C|A) + 7I(A;B|D) + 7I(C;D) \\
&\ldots
\end{aligned}$$



$$I(A;B) \leq (x+1)I(A;B|C) + yI(A;C|B) + zI(B;C|A) + I(A;B|D) + I(C;D)$$

Do these methods suffice (in the limit) to generate all (four-variable) information inequalities?

The four-atom conjecture

The extreme rays of $\overline{\Gamma_4^*}$ at or near the minus Ingleton direction are attained by probability distributions where the four random variables are over the alphabet $\{0, 1\}$ and the only 4-tuples with nonzero probability are 0000, 0011, 0101, and 1010.

Linear rank inequalities

Any inequality which always holds for entropies will always hold for ranks of linear subspaces; however, there are inequalities that always hold for ranks of linear subspaces but sometimes fail for entropies.

The Ingleton inequality

$$I(A;B) \leq I(A;B|C) + I(A;B|D) + I(C;D)$$

This inequality and the Shannon inequalities generate all linear rank inequalities on four variables. (Hammer-Romashchenko-Shen-Vereshchagin)

Common informations

Random variable Z is a *common information* of random variables A and B if:

$$H(Z|A) = 0$$
$$H(Z|B) = 0$$
$$H(Z) = I(A; B)$$

Common informations do not always exist, but they do in the case that A and B come from vector subspaces (Z will correspond to the intersection of these subspaces).

Proof of the Ingleton inequality

Lemma:

$$I(A;B|C) + H(Z|A) \geq I(Z;B|C)$$

Proof of Ingleton:

$$\begin{aligned}
I(A;B|C)+I(A;B|D) + I(C;D) + 2H(Z|A) + 2H(Z|B) \\
&\geq I(A;Z|C) + I(A;Z|D) + I(C;D) + 2H(Z|A) \\
&\geq I(Z;Z|C) + I(Z;Z|D) + I(C;D) \\
&= H(Z|C) + H(Z|D) + I(C;D) \\
&\geq H(Z|C) + I(C;Z) \\
&\geq I(Z;Z) \\
&= H(Z).
\end{aligned}$$

Now let Z be a common information for A and B.

Procedure for generating inequalities

- List known inequalities
- Determine extreme rays of polytope
- Discard rays known to be representable
- List common information specifications
- Test extreme rays against common information specifications
- If a test fails, determine new inequality and start over

New five-variable linear rank inequalities

$$I(A;B) \leq I(A;B|C) + I(A;B|D) + I(C;D|E) + I(A;E)$$

$$I(A;B) \leq I(A;B|C) + I(A;C|D) + I(A;D|E) + I(B;E)$$

$$I(A;B) \leq I(A;C) + I(A;B|D) + I(B;E|C) + I(A;D|CE)$$

$$\dots$$

$$\begin{aligned} I(A;CD) + I(B;CD) \leq\ & I(B;D) + I(B;C|E) + I(C;E|D) + I(A;E) \\ & + I(A;C|BD) + I(AB;D|C) + I(A;D|BE) + I(A;B|DE) \end{aligned}$$

Extreme rays for five-variable linear ranks

162 extreme rays (7943 counting permutations)

All have been verified to be representable over any sufficiently large field.

Six or more variables

Work in progress.

The End.