Entropic Inequalities and Marginal Problems*

Jayant Apte
ASPITRG
Drexel University

*FRITZ, T., & CHAVES, R. (2013). Entropic Inequalities and Marginal Problems. IEEE transactions on information theory, 59(2), 803-817.

The three coins

- You have 3 coins A_1, A_2 and A_3
- You can flip only two of them at once

The three coins

A_1	A_2	P
T	T	0.5
T	H	0
H	T	0
H	H	0.5

A_1	A_3	P
T	T	0
T	H	0.5
H	T	0.5
H	H	0

A_2	A_3	P
T	T	0.5
T	H	0
H	T	0
H	H	0.5

The three coins

A_1	A_2	P
T	T	0.5
T	H	0
H	T	0
H	H	0.5

Joint Distribution on A_1 and A_2

A_2	A_3	P
T	T	0.5
T	H	0
H	T	0
H	H	0.5

Joint Distribution on A_2 and A_3

A_1	A_3	P
T	T	0
T	H	0.5
H	T	0.5
H	H	0

Joint Distribution on A_1 and A_3

The three coins

Joint distribution on A_1, A_2, A_3?

An attempt

A_1 and A_2 are perfectly correlated

An attempt

A_1 and A_2 are perfectly correlated

A_2 and A_3 are perfectly correlated

An attempt

A_1 and A_2 are perfectly correlated
A_2 and A_3 are perfectly correlated

$\searrow$

A_1 and A_3 should be perfectly correlated

An attempt

A_1 and A_2 are perfectly correlated
A_2 and A_3 are perfectly correlated

↘

A_1 and A_3 should be perfectly correlated

But they are not!!!

Conclusion

- There is no joint distribution compatible with given marginals

Conclusion

- There is no joint distribution compatible with given marginals
- There is some systematic error in the coin flips

Conclusion

- There is no joint distribution compatible with given marginals
- There is some systematic error in the coin flips
- OR The coins are operated by some weird mechanism creating the required distribution as a function depending on which two are flipped together

Outline

- Marginal scenario and marginal model
- Contextuality
- Marginal problem
- Partial polymatroids
- Relationship between marginal models and partial polymatroids
- Marginal Problem for polymatroids
- Computations
- Inference of common ancestors in Bayesian Networks
- Contexuality of the n-Cycle marginal scenario
- Extention to continuous random variables

Notation

- $[n] = \{1, \ldots, n\}$
- $2^{[n]}$= The set of all subsets of $[n]$
- Entropy is in bits
- All random variables are discrete
- For any subset $S \in [n]$, $A_S = (A_i)_{i \in S}$
- $A_{[n]}$ is the joint distribution of all variables
- If P is probability distribution of some superset of $S \subseteq \mathcal{M}$ then $P_{|S}$ is the marginal distribution associated with A_S

Marginal Scenario

- A *marginal scenario* $\mathcal{M}$ on $[n]$ is a nonempty collection $\mathcal{M} = \{S_1, \ldots, S_{|\mathcal{M}|}\}$ of subsets $S_i \subseteq [n]$ such that if $S \in \mathcal{M}$ and $S^{'} \subseteq S$, then also $S^{'} \in \mathcal{M}$

- It suffices to only specify all inclusionwise maximal subsets called *measurement contexts*

- A marginal scenario is independent sets of a matroid except that it need not satisfy the augmentation axiom

- For a set system $\mathcal{X} \in 2^{[n]}$, $\bar{X}^{\subseteq}$ is set system $\mathcal{X}$ along with all subsets of $\mathcal{X}$

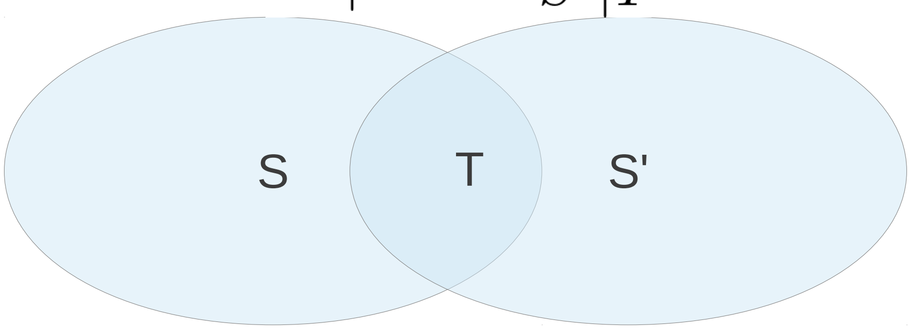
Marginal Model

- Denote by $P^{\mathcal{M}}$ a collection of probability distributions on subsets of $\mathcal{M}$
- $P^{\mathcal{M}}$ is a *marginal model* if it satisfies the condition:

  For any pair $S, T \in \mathcal{M}$ with $T \subseteq S$, $P^{\mathcal{M}}_{S|T} = P^{\mathcal{M}}_{T}$

- A trivial marginal model can be obtained by starting with a joint distribution and taking its marginals
- An implication of compatibility condition:

$$P^{\mathcal{M}}_{S|T} = P^{\mathcal{M}}_{S'|T}$$



Contexuality

- $P^{\mathcal{M}}$ is *noncontexual* if there exists a joint distribution $P = P(a_1, \ldots, a_n)$ such that its marginals coincide with the distributions occuring in the marginal model
- Otherwise $P^{\mathcal{M}}$ is *contexual*

Example

- Formalization of the "three-coins" example
- Consider a marginal scenario $\mathcal{M} = \mathcal{C}_3$
  $\mathcal{C}_3 = \overline{\{\{1,2\},\{1,3\},\{2,3\}\}}^{\subseteq}$
- The two variable distributions are:

$$P^{\mathcal{C}_3}_{\{1,2\}}(A_1 = a_1, A_2 = a_2) = \begin{cases} 1/2 \text{ if } a_1 = a_2 \\ 0 \text{ if } a_1 \neq a_2 \end{cases}$$

$$P^{\mathcal{C}_3}_{\{2,3\}}(A_2 = a_2, A_3 = a_3) = \begin{cases} 1/2 \text{ if } a_2 = a_3 \\ 0 \text{ if } a_2 \neq a_3 \end{cases}$$

$$P^{\mathcal{C}_3}_{\{1,3\}}(A_1 = a_1, A_3 = a_3) = \begin{cases} 0 \text{ if } a_1 = a_3 \\ 1/2 \text{ if } a_1 \neq a_3 \end{cases}$$

- $P^{\mathcal{C}_3}_{\{1\}}, P^{\mathcal{C}_3}_{\{2\}}$ and $P^{\mathcal{C}_3}_{\{3\}}$ can be determined from above

Marginal problem

How to decide if system of distributions $P^{\mathcal{C}_3}$ is contexual or non-contexual?

A naive probability space approach

- Consider n discrete random variables $X_1, \ldots, X_n$, each with d outcomes
- We are given a distribution system $P^{\mathcal{M}}$ corresponding to some marginal scenario $\mathcal{M}$
- We set up an LP in $\mathbb{R}^{d^n}$
- We constrain this LP by following constraints:
  - $x_1 + \ldots + x_{d^n} = 1, x_i \geq 0 \;\; \forall i \in [d^n]$
  - Every subset of $\mathcal{M}$ of size k gives d^k constraints from the marginal on that subset
- Infeasible LP $\Rightarrow P^{\mathcal{M}}$ is contexual
- Feasible LP $\Rightarrow P^{\mathcal{M}}$ is non- contexual

Example: 3 coins

- We setup LP in $\mathbb{R}^7$. (since there are 7 outcomes: TTT,TTH,THT,THH,HTT,HTH,HHT,HHH)
- $\sum_{i=1}^{7} x_i = 1$
- $x_i \geq 0 \;\; \forall i \in \{1, \ldots, 7\}$
- The marginal constraints are of type:

$$\begin{aligned} p(A_1 = T, A_2 = T) =& p(A_1 = T, A_2 = T, A_3 = H) + p(A_1 = T, A_2 = T, A_3 = H) \\ =& 1/2 \\ \therefore x_2 + x_1 =& 1/2 \end{aligned}$$

Example: 3 coins

- We setup LP in $\mathbb{R}^7$. (since there are 7 outcomes: TTT,TTH,THT,THH,HTT,HTH,HHT,HHH)
- $\sum_{i=1}^{7} x_i = 1$
- $x_i \geq 0 \ \ \forall i \in \{1, \ldots, 7\}$
- The marginal constraints are of type:

$$\begin{aligned} p(A_1 = T, A_2 = T) =& p(A_1 = T, A_2 = T, A_3 = H) + p(A_1 = T, A_2 = T, A_3 = H) \\ =& 1/2 \\ \therefore x_2 + x_1 =& 1/2 \end{aligned}$$

This LP turns out to be infeasible

The entropy cone

- Consider n random variables $A_1, \ldots, A_n$ with some joint distribution P
- For any $S \subseteq [n]$,

$$H(A_S) = -\sum_{a_S} P_{|S}(a_S) \log P_{|S}(a_S)$$

- By convention $H(A_\phi) = 0$
- The vector $(H(A_\phi), \ldots, H(A_{[n]}))$ is a point in $\mathbb{R}^{2^{[n]}}$
- The closure of set of all such points is called the *entropy cone* $\overline{\Gamma}_n^*$

Basic inequalities

- There are some of the obvious constraints satisfied by $\overline{\Gamma}^*_N$

$$H(A_S) \geq 0 \quad (1)$$

$$H(A_T) \leq H(A_S) \ \ \forall T \subseteq S \quad (2)$$

$$H(A_S) + H(A_T) \geq H(A_{S \cap T}) + H(A_{S \cup T}) \quad (3)$$

- An inequality id called *shannon type* if it can be written as non-negative linear combination of basic inequalities

Polymatroids

- *Polymatroid* is a pair $([n], f)$ where $n \in \mathbb{N}$ and f is the *rank function*, a function $f : 2^{[n]} \mapsto \mathbb{R}$ s.t. f satisfies all the basic inequalities

$$f(S) \geq 0 \tag{1}$$
$$f(T) \leq f(S) \quad \forall T \subseteq S \tag{2}$$
$$f(S) + f(T) \geq f(S \cap T) + f(S \cup T) \tag{3}$$

- $([n], H)$ where H is the entropy function is a polymatroid

Partial Polymatroids

- A partial polymatroid $f^{\mathcal{M}}$ on a marginal scenario $\mathcal{M}$ is a function
$$f^{\mathcal{M}} : \mathcal{M} \mapsto \mathbb{R}$$
which satisfies basic inequalities for all $S, T \subseteq \mathcal{M}$ for which $S \cup T \in \mathcal{M}$

  - A trivial way of obtaining partial polymatroid is to restrict a polymatroid to subsets in $\mathcal{M}$

  - Unfortunately this not the only way partial polymatroids arise

Contexuality of partial polymatroids

- A partial polymatroid $f^{\mathcal{M}}$ is *noncontexual* if there is a polymatroid f such that $f^{\mathcal{M}}(S) = f(S)$ for all $S \in \mathcal{M}$. Otherwise $f^{\mathcal{M}}$ is contexual.
- *Marginal Problem for Polymatroids*: Given a partial polymatroid $f^{\mathcal{M}}$ on $\mathcal{M}$, under which conditions it is non-contexual?

Example: Partial polymatroids over $\mathcal{C}_3$

- Consider the three coins scenario: $\mathcal{C}_3 = \overline{\{\{1,2\},\{1,3\},\{2,3\}\}}^{\subseteq}$

- Two basic inequalities for any polymatroid $(f,[3])$ are:

$$f(\{1,3\}) \leq f(\{1,2,3\}) \quad (1)$$

$$f(\{1,2,3\}) + f(\{2\}) \leq f(\{1,2\}) + f(\{2,3\}) \quad (2)$$

- Adding (1) and (2) we get:

$$f(\{1,3\}) + f(\{2\}) \leq f(\{1,2\}) + f(\{2,3\})$$

Triangle inequality

Example: Modified 3-coins

- The two variable distributions are:

$$P^{\mathcal{C}_3}_{\{1,2\}}(A_1 = a_1, A_2 = a_2) = \begin{cases} 1/2 \text{ if } a_1 = a_2 \\ 0 \text{ if } a_1 \neq a_2 \end{cases}$$

$$P^{\mathcal{C}_3}_{\{2,3\}}(A_2 = a_2, A_3 = a_3) = \begin{cases} 1/2 \text{ if } a_2 = a_3 \\ 0 \text{ if } a_2 \neq a_3 \end{cases}$$

$$P^{\mathcal{C}_3}_{\{1,3\}}(A_1 = a_1, A_3 = a_3) = 1/4 \ \ \forall a_1, a_3$$

- $P^{\mathcal{C}_3}_{\{1\}}, P^{\mathcal{C}_3}_{\{2\}}$ and $P^{\mathcal{C}_3}_{\{3\}}$ can be determined from above
- $f^{\mathcal{C}_3}(\{1\}) = f^{\mathcal{C}_3}(\{2\}) = f(\{3\}) = 1,$
  $f^{\mathcal{C}_3}(\{1,2\}) = f^{\mathcal{C}_3}(\{2,3\}) = 1,$
  $f^{\mathcal{C}_3}(\{1,2\}) = 2$
- Clearly, the triangle inequality is violated

References

- FRITZ, T., & CHAVES, R. (2013). Entropic Inequalities and Marginal Problems. IEEE transactions on information theory, 59(2), 803-817.