

Chapter 3 The *I*-Measure

© Raymond W. Yeung 2014 The Chinese University of Hong Kong

• Set-theoretic structure of Shannon's information measures

- *•* Set-theoretic structure of Shannon's information measures
- *•* Why Shannon's information measures can be represented by a form of Venn diagram – information diagram?

- *•* Set-theoretic structure of Shannon's information measures
- Why Shannon's information measures can be represented by a form of Venn diagram – information diagram?
- *•* How to use information diagrams to obtain information identities and inequalities

- *•* Set-theoretic structure of Shannon's information measures
- Why Shannon's information measures can be represented by a form of Venn diagram – information diagram?
- *•* How to use information diagrams to obtain information identities and inequalities
- Problem-solving examples

$$
\begin{array}{cccc}\n\bigoplus \text{ } & H/I & \leftrightarrow & \mu^* \\
\bigoplus \text{ } & \text{ } & \text{ } & \text{ } & \\
 & \vdots & \leftrightarrow & \text{ } & \\
 & \vdots & \leftrightarrow & \text{ } & \\
 & \downarrow & \leftrightarrow & \text{ } & \text{ } & \\
 & \downarrow & \leftrightarrow & \text{ } & \text{ } & \\
 & \downarrow & \leftrightarrow & \text{ } & \text{ } & \\
 & \downarrow & \leftrightarrow & \text{ } & \text{ } & \\
 & \downarrow & \leftrightarrow & \text{ } & \text{ } & & \\
 & \downarrow & \leftrightarrow & \text{ } & \text{ } & & \\
 & \downarrow & \leftrightarrow & \text{ } & \text{ } & & \\
 & \downarrow & \leftrightarrow & \text{ } & \text{ } & & \\
 & \downarrow & \leftrightarrow & \text{ } & \text{ } & & & \\
 & \downarrow & \leftrightarrow & \text{ } & \text{ } & & & \\
 & \downarrow & \leftrightarrow & \text{ } & \text{ } & & & \\
 & \downarrow & \leftrightarrow & \text{ } & \text{ } & & & \\
 & \downarrow & \leftrightarrow & \text{ } & \text{ } & & & \\
 & \downarrow & \leftrightarrow & \text{ } & & & & \\
 & \downarrow & \leftrightarrow & \text{ } & & & & \\
 & \downarrow & \leftrightarrow & \text{ } & & & & \\
 & \downarrow & \leftrightarrow & \text{ } & & & & \\
 & \downarrow & \leftrightarrow & \text{ } & & & & \\
 & \downarrow & \leftrightarrow & \text{ } & & & & \\
 & \downarrow & \leftrightarrow & \text{ } & & & & \\
 & \downarrow & \leftrightarrow & \
$$

$$
\begin{array}{cccc}\n & H/I & \leftrightarrow & \mu^* \\
 & \downarrow & & \downarrow \\
\text{E}\rightarrow & & \downarrow & & \uparrow \\
 & & \downarrow & & \rightarrow & \left(A - B = A \cap B^c \right)\n\end{array}
$$

$$
H/I \leftrightarrow \mu^*
$$
\n
$$
\leftrightarrow \cup
$$
\n
$$
\Leftrightarrow \land \cup
$$
\n
$$
\Box
$$
\n
$$
\Leftrightarrow \land \cup
$$
\n
$$
(A - B = A \cap B^c)
$$

$$
H/I \leftrightarrow \mu^*
$$

, $\leftrightarrow \cup$
; $\leftrightarrow \cap$
, $\leftrightarrow \cap$
, $\leftrightarrow \wedge$
, $\Rightarrow (A - B = A \cap B^c)$

- μ^* is some signed measure (set-additive function).
- *•* Examples:

$$
H(X_1|X_2) = \mu^*(\tilde{X}_1 - \tilde{X}_2)
$$

\n
$$
H(X_2|X_1) = \mu^*(\tilde{X}_2 - \tilde{X}_1)
$$

\n
$$
I(X_1; X_2) = \mu^*(\tilde{X}_1 \cap \tilde{X}_2).
$$

$$
H/I \leftrightarrow \mu^* \n, \leftrightarrow \cup \n; \leftrightarrow \cap \n\downarrow \leftrightarrow - (A - B = A \cap B^c)
$$

- μ^* is some signed measure (set-additive function).
- *•* Examples:

$$
\begin{array}{rcl}\n\bigoplus & H(X_1|X_2) & = & \mu^*(\tilde{X}_1 - \tilde{X}_2) \\
H(X_2|X_1) & = & \mu^*(\tilde{X}_2 - \tilde{X}_1) \\
I(X_1; X_2) & = & \mu^*(\tilde{X}_1 \cap \tilde{X}_2).\n\end{array}
$$

$$
H/I \leftrightarrow \mu^* \n, \leftrightarrow \cup \n; \leftrightarrow \cap \n\downarrow \leftrightarrow -(A - B = A \cap B^c)
$$

- μ^* is some signed measure (set-additive function).
- *•* Examples:

$$
H(X_1|X_2) = \mu^*(\tilde{X}_1 - \tilde{X}_2)
$$

\n
$$
H(X_2|X_1) = \mu^*(\tilde{X}_2 - \tilde{X}_1)
$$

\n
$$
I(X_1; X_2) = \mu^*(\tilde{X}_1 \cap \tilde{X}_2).
$$

$$
H/I \leftrightarrow \mu^* \n, \leftrightarrow \cup \n; \leftrightarrow \cap \n| \leftrightarrow - (A - B = A \cap B^c)
$$

- μ^* is some signed measure (set-additive function).
- *•* Examples:

$$
H(X_1|X_2) = \mu^*(\tilde{X}_1 - \tilde{X}_2)
$$

\n
$$
H(X_2|X_1) = \mu^*(\tilde{X}_2 - \tilde{X}_1)
$$

\n
$$
I(X_1; X_2) = \mu^*(\tilde{X}_1 \cap \tilde{X}_2).
$$

$$
\mu^*(\tilde{X}_1 \cup \tilde{X}_2) = \mu^*(\tilde{X}_1) + \mu^*(\tilde{X}_2) - \mu^*(\tilde{X}_1 \cap \tilde{X}_2)
$$

corresponds to

$$
H(X_1, X_2) = H(X_1) + H(X_2) - I(X_1; X_2)
$$

$$
\mu^*(\tilde{X}_1 \cup \tilde{X}_2) = \mu^*(\tilde{X}_1) + \mu^*(\tilde{X}_2) - \mu^*(\tilde{X}_1 \cap \tilde{X}_2)
$$

corresponds to

$$
H(X_1, X_2) = H(X_1) + H(X_2) - I(X_1; X_2)
$$

$$
\mu^*(\tilde{X}_1 \cup \tilde{X}_2) = \mu^*(\tilde{X}_1) + \mu^*(\tilde{X}_2) - \mu^*(\tilde{X}_1 \cap \tilde{X}_2)
$$

corresponds to

$$
H(\underline{X_1}, X_2) = H(X_1) + H(X_2) - I(X_1; X_2)
$$

$$
\mu^*(\tilde{X}_1 \cup \tilde{X}_2) = \mu^*(\tilde{X}_1) + \mu^*(\tilde{X}_2) - \mu^*(\tilde{X}_1 \cap \tilde{X}_2)
$$

corresponds to

$$
H(X_1, X_2) = H(X_1) + H(X_2) - I(X_1; X_2)
$$

$$
\mu^*(\tilde{X}_1 \cup \tilde{X}_2) = \mu^*(\tilde{X}_1) + \mu^*(\tilde{X}_2) - \mu^*(\tilde{X}_1 \cap \tilde{X}_2)
$$

corresponds to

$$
H(X_1, \underline{X_2}) = H(X_1) + H(X_2) - I(X_1; X_2)
$$

$$
\mu^*(\tilde{X}_1 \cup \tilde{X}_2) = \underline{\mu^*}(\tilde{X}_1) + \mu^*(\tilde{X}_2) - \mu^*(\tilde{X}_1 \cap \tilde{X}_2)
$$

corresponds to

$$
H(X_1, X_2) = \underline{H}(X_1) + H(X_2) - I(X_1; X_2)
$$

$$
\mu^*(\tilde{X}_1 \cup \tilde{X}_2) = \mu^*(\underline{\tilde{X}_1}) + \mu^*(\tilde{X}_2) - \mu^*(\tilde{X}_1 \cap \tilde{X}_2)
$$

corresponds to

$$
H(X_1, X_2) = H(\underline{X_1}) + H(X_2) - I(X_1; X_2)
$$

$$
\mu^*(\tilde{X}_1 \cup \tilde{X}_2) = \mu^*(\tilde{X}_1) + \underline{\mu^*}(\tilde{X}_2) - \mu^*(\tilde{X}_1 \cap \tilde{X}_2)
$$

corresponds to

$$
H(X_1, X_2) = H(X_1) + \underline{H}(X_2) - I(X_1; X_2)
$$

$$
\mu^*(\tilde{X}_1 \cup \tilde{X}_2) = \mu^*(\tilde{X}_1) + \mu^*(\tilde{\underline{X}}_2) - \mu^*(\tilde{X}_1 \cap \tilde{X}_2)
$$

corresponds to

$$
H(X_1, X_2) = H(X_1) + H(\underline{X_2}) - I(X_1; X_2)
$$

$$
\mu^*(\tilde{X}_1 \cup \tilde{X}_2) = \mu^*(\tilde{X}_1) + \mu^*(\tilde{X}_2) - \mu^*(\tilde{X}_1 \cap \tilde{X}_2)
$$

corresponds to

$$
H(X_1, X_2) = H(X_1) + H(X_2) - \underline{I}(X_1; X_2)
$$

$$
\mu^*(\tilde{X}_1 \cup \tilde{X}_2) = \mu^*(\tilde{X}_1) + \mu^*(\tilde{X}_2) - \mu^*(\underline{\tilde{X}}_1 \cap \tilde{X}_2)
$$

corresponds to

$$
H(X_1, X_2) = H(X_1) + H(X_2) - I(\underline{X_1}; X_2)
$$

$$
\mu^*(\tilde{X}_1 \cup \tilde{X}_2) = \mu^*(\tilde{X}_1) + \mu^*(\tilde{X}_2) - \mu^*(\tilde{X}_1 \cap \tilde{X}_2)
$$

corresponds to

$$
H(X_1, X_2) = H(X_1) + H(X_2) - I(X_1; X_2)
$$

$$
\mu^*(\tilde{X}_1 \cup \tilde{X}_2) = \mu^*(\tilde{X}_1) + \mu^*(\tilde{X}_2) - \mu^*(\tilde{X}_1 \cap \tilde{X}_2)
$$

corresponds to

$$
H(X_1, X_2) = H(X_1) + H(X_2) - I(X_1; \underline{X_2})
$$

3.1 Preliminaries

Definition 3.1 The field \mathcal{F}_n generated by sets $\tilde{X}_1, \tilde{X}_2, \cdots, \tilde{X}_n$ is the collection of sets which can be obtained by any sequence of usual set operations (union, intersection, complement, and difference) on $\tilde{X}_1, \tilde{X}_2, \cdots, \tilde{X}_n$.

Definition 3.1 The field \mathcal{F}_n generated by sets $\tilde{X}_1, \tilde{X}_2, \cdots, \tilde{X}_n$ is the collection of sets which can be obtained by any sequence of usual set operations (union, intersection, complement, and difference) on $\tilde{X}_1, \tilde{X}_2, \cdots, \tilde{X}_n$.

Definition 3.2 The atoms of \mathcal{F}_n are sets of the form $\bigcap_{i=1}^n Y_i$, where Y_i is either \tilde{X}_i or \tilde{X}_i^c , the complement of \tilde{X}_i .

• The sets \tilde{X}_1 and \tilde{X}_2 generate the field \mathcal{F}_2 .

- The sets \tilde{X}_1 and \tilde{X}_2 generate the field \mathcal{F}_2 .
- There are 4 atoms in \mathcal{F}_2 :

$$
\tilde{X}_1 \cap \tilde{X}_2, \ \tilde{X}_1^c \cap \tilde{X}_2, \ \tilde{X}_1 \cap \tilde{X}_2^c, \ \tilde{X}_1^c \cap \tilde{X}_2^c
$$

- The sets \tilde{X}_1 and \tilde{X}_2 generate the field \mathcal{F}_2 .
- There are 4 atoms in \mathcal{F}_2 :

$$
\tilde{X}_1 \cap \tilde{X}_2, \ \tilde{X}_1^c \cap \tilde{X}_2, \ \tilde{X}_1 \cap \tilde{X}_2^c, \ \tilde{X}_1^c \cap \tilde{X}_2^c
$$

- The sets \tilde{X}_1 and \tilde{X}_2 generate the field \mathcal{F}_2 .
- There are 4 atoms in \mathcal{F}_2 :

$$
\tilde{X}_1 \cap \tilde{X}_2, \ \underline{\tilde{X}_1^c} \cap \tilde{X}_2, \ \tilde{X}_1 \cap \tilde{X}_2^c, \ \tilde{X}_1^c \cap \tilde{X}_2^c
$$

- The sets \tilde{X}_1 and \tilde{X}_2 generate the field \mathcal{F}_2 .
- There are 4 atoms in \mathcal{F}_2 :

$$
\tilde{X}_1 \cap \tilde{X}_2, \ \tilde{X}_1^c \cap \tilde{X}_2, \ \tilde{\underline{X}}_1 \cap \tilde{X}_2^c, \ \tilde{X}_1^c \cap \tilde{X}_2^c
$$

- The sets \tilde{X}_1 and \tilde{X}_2 generate the field \mathcal{F}_2 .
- There are 4 atoms in \mathcal{F}_2 :

$$
\tilde{X}_1 \cap \tilde{X}_2, \ \tilde{X}_1^c \cap \tilde{X}_2, \ \tilde{X}_1 \cap \tilde{X}_2^c, \ \underline{\tilde{X}_1^c} \cap \tilde{X}_2^c
$$

- The sets \tilde{X}_1 and \tilde{X}_2 generate the field \mathcal{F}_2 .
- There are 4 atoms in \mathcal{F}_2 :

$$
\tilde{X}_1 \cap \tilde{X}_2, \ \tilde{X}_1^c \cap \tilde{X}_2, \ \tilde{X}_1 \cap \tilde{X}_2^c, \ \tilde{X}_1^c \cap \tilde{X}_2^c
$$

• There are a total of $2^4 = 16$ sets in \mathcal{F}_2 , formed by the unions of the above 4 atoms.

- The sets \tilde{X}_1 and \tilde{X}_2 generate the field \mathcal{F}_2 .
- There are 4 atoms in \mathcal{F}_2 :

$$
\tilde{X}_1 \cap \tilde{X}_2, \ \tilde{X}_1^c \cap \tilde{X}_2, \ \tilde{X}_1 \cap \tilde{X}_2^c, \ \tilde{X}_1^c \cap \tilde{X}_2^c
$$

 $2^{\text{-}} = 10$ • There are a total of $2^4 = 16$ sets in \mathcal{F}_2 , formed by the unions of the above 4 atoms.

 $\mu(A \cup B) = \mu(A) + \mu(B).$

 $\mu(\underline{A} \cup B) = \mu(\underline{A}) + \mu(B).$

 $\mu(A \cup B) = \mu(A) + \mu(B).$

$$
\mu(A \cup B) = \mu(A) + \mu(B).
$$

Remark For any set-additive function *µ*,

$$
\mu(\emptyset)=0
$$

because for any set *A*,

$$
\mu(A) = \mu(A \cup \emptyset) = \mu(A) + \mu(\emptyset)
$$

$$
\mu(A \cup B) = \mu(A) + \mu(B).
$$

Remark For any set-additive function *µ*,

$$
\mu(\emptyset)=0
$$

because for any set *A*,

$$
\mu(\underline{A}) = \mu(A \cup \emptyset) = \mu(A) + \mu(\emptyset)
$$

$$
\mu(A \cup B) = \mu(A) + \mu(B).
$$

Remark For any set-additive function *µ*,

$$
\mu(\emptyset)=0
$$

because for any set *A*,

$$
\mu(\underline{A}) = \mu(\underline{A \cup \emptyset}) = \mu(A) + \mu(\emptyset)
$$

$$
\mu(A \cup B) = \mu(A) + \mu(B).
$$

Remark For any set-additive function *µ*,

$$
\mu(\emptyset)=0
$$

because for any set *A*,

$$
\mu(A) = \mu(A \cup \emptyset) = \underline{\mu(A)} + \mu(\emptyset)
$$

$$
\mu(A \cup B) = \mu(A) + \mu(B).
$$

Remark For any set-additive function *µ*,

$$
\mu(\emptyset)=0
$$

because for any set *A*,

$$
\mu(A) = \mu(A \cup \emptyset) = \mu(A) + \mu(\emptyset)
$$

$$
\mu(A \cup B) = \mu(A) + \mu(B).
$$

Remark For any set-additive function *µ*,

$$
\mu(\emptyset)=0
$$

because for any set *A*,

$$
\mu(A) = \mu(A \cup \emptyset) = \mu(A) + \mu(\emptyset)
$$

$$
\mu(A \cup B) = \mu(A) + \mu(B).
$$

Remark For any set-additive function *µ*,

 $\mu(\emptyset)=0$

because for any set *A*,

$$
\mu(A) = \mu(A \cup \emptyset) = \mu(A) + \mu(\emptyset)
$$

$$
\mu(A \cup B) = \mu(A) + \mu(B).
$$

Remark For any set-additive function *µ*,

 $\mu(\emptyset)=0$

because for any set *A*,

$$
\mu(A) = \mu(A \cup \emptyset) = \mu(A) + \mu(\emptyset)
$$

$$
\mu(A \cup B) = \mu(A) + \mu(B).
$$

Remark For any set-additive function *µ*,

$$
\mu(\emptyset)=0
$$

because for any set *A*,

$$
\mu(A) = \mu(A \cup \emptyset) = \mu(A) + \mu(\emptyset)
$$

implies $\mu(\emptyset) = 0$.

Remark A signed measure can take positive or negative values. If a signed measure takes only positive values, it is simply called a measure.

• A signed measure μ on \mathcal{F}_2 is completely specified by the values on the atoms

 $\mu(\tilde{X}_1 \cap \tilde{X}_2), \ \mu(\tilde{X}_1^c \cap \tilde{X}_2), \ \mu(\tilde{X}_1 \cap \tilde{X}_2^c), \ \mu(\tilde{X}_1^c \cap \tilde{X}_2^c)$

• A signed measure μ on \mathcal{F}_2 is completely specified by the values on the atoms

 $\mu(\tilde{X}_1 \cap \tilde{X}_2), \ \mu(\tilde{X}_1^c \cap \tilde{X}_2), \ \mu(\tilde{X}_1 \cap \tilde{X}_2^c), \ \mu(\tilde{X}_1^c \cap \tilde{X}_2^c)$

- A signed measure μ on \mathcal{F}_2 is completely specified by the values on the atoms $\mu(\tilde{X}_1 \cap \tilde{X}_2), \ \mu(\tilde{X}_1^c \cap \tilde{X}_2), \ \mu(\tilde{X}_1 \cap \tilde{X}_2^c), \ \mu(\tilde{X}_1^c \cap \tilde{X}_2^c)$
- The value of μ on other sets in \mathcal{F}_2 are obtained by set-additivity. For example,

$$
\mu(\tilde{X}_1) = \mu((\tilde{X}_1 \cap \tilde{X}_2) \cup (\tilde{X}_1 \cap \tilde{X}_2)) \n= \mu(\tilde{X}_1 \cap \tilde{X}_2) + \mu(\tilde{X}_1 \cap \tilde{X}_2^c)
$$

- A signed measure μ on \mathcal{F}_2 is completely specified by the values on the atoms $\mu(\tilde{X}_1 \cap \tilde{X}_2), \ \mu(\tilde{X}_1^c \cap \tilde{X}_2), \ \mu(\tilde{X}_1 \cap \tilde{X}_2^c), \ \mu(\tilde{X}_1^c \cap \tilde{X}_2^c)$
- The value of μ on other sets in \mathcal{F}_2 are obtained by set-additivity. For example,

$$
\mu(\tilde{X}_1) = \mu((\tilde{X}_1 \cap \tilde{X}_2) \cup (\tilde{X}_1 \cap \tilde{X}_2))
$$

=
$$
\mu(\tilde{X}_1 \cap \tilde{X}_2) + \mu(\tilde{X}_1 \cap \tilde{X}_2^c)
$$

- A signed measure μ on \mathcal{F}_2 is completely specified by the values on the atoms $\mu(\tilde{X}_1 \cap \tilde{X}_2), \ \mu(\tilde{X}_1^c \cap \tilde{X}_2), \ \mu(\tilde{X}_1 \cap \tilde{X}_2^c), \ \mu(\tilde{X}_1^c \cap \tilde{X}_2^c)$
- The value of μ on other sets in \mathcal{F}_2 are obtained by set-additivity. For example,

$$
\mu(\tilde{X}_1) = \mu((\tilde{X}_1 \cap \tilde{X}_2) \cup (\tilde{X}_1 \cap \tilde{X}_2))
$$

=
$$
\mu(\tilde{X}_1 \cap \tilde{X}_2) + \mu(\tilde{X}_1 \cap \tilde{X}_2^c)
$$

- A signed measure μ on \mathcal{F}_2 is completely specified by the values on the atoms $\mu(\tilde{X}_1 \cap \tilde{X}_2), \ \mu(\tilde{X}_1^c \cap \tilde{X}_2), \ \mu(\tilde{X}_1 \cap \tilde{X}_2^c), \ \mu(\tilde{X}_1^c \cap \tilde{X}_2^c)$
- The value of μ on other sets in \mathcal{F}_2 are obtained by set-additivity. For example,

$$
\mu(\tilde{X}_1) = \mu((\tilde{X}_1 \cap \tilde{X}_2) \cup (\tilde{X}_1 \cap \tilde{X}_2))
$$

=
$$
\mu(\tilde{X}_1 \cap \tilde{X}_2) + \mu(\tilde{X}_1 \cap \tilde{X}_2^c)
$$

- A signed measure μ on \mathcal{F}_2 is completely specified by the values on the atoms $\mu(\tilde{X}_1 \cap \tilde{X}_2), \ \mu(\tilde{X}_1^c \cap \tilde{X}_2), \ \mu(\tilde{X}_1 \cap \tilde{X}_2^c), \ \mu(\tilde{X}_1^c \cap \tilde{X}_2^c)$
- The value of μ on other sets in \mathcal{F}_2 are obtained by set-additivity. For example,

$$
\mu(\tilde{X}_1) = \mu((\tilde{X}_1 \cap \tilde{X}_2) \cup (\tilde{X}_1 \cap \tilde{X}_2)) \n= \mu(\tilde{X}_1 \cap \tilde{X}_2) + \mu(\tilde{X}_1 \cap \tilde{X}_2^c)
$$

- A signed measure μ on \mathcal{F}_2 is completely specified by the values on the atoms $\mu(\tilde{X}_1 \cap \tilde{X}_2), \ \mu(\tilde{X}_1^c \cap \tilde{X}_2), \ \mu(\tilde{X}_1 \cap \tilde{X}_2^c), \ \mu(\tilde{X}_1^c \cap \tilde{X}_2^c)$
- The value of μ on other sets in \mathcal{F}_2 are obtained by set-additivity. For example,

$$
\mu(\tilde{X}_1) = \mu((\tilde{X}_1 \cap \tilde{X}_2) \cup (\tilde{X}_1 \cap \tilde{X}_2))
$$

=
$$
\mu(\tilde{X}_1 \cap \tilde{X}_2) + \mu(\tilde{X}_1 \cap \tilde{X}_2^c)
$$

- A signed measure μ on \mathcal{F}_2 is completely specified by the values on the atoms $\mu(\tilde{X}_1 \cap \tilde{X}_2), \ \mu(\tilde{X}_1^c \cap \tilde{X}_2), \ \mu(\tilde{X}_1 \cap \tilde{X}_2^c), \ \mu(\tilde{X}_1^c \cap \tilde{X}_2^c)$
- The value of μ on other sets in \mathcal{F}_2 are obtained by set-additivity. For example,

$$
\mu(\tilde{X}_1) = \mu((\tilde{X}_1 \cap \tilde{X}_2) \cup (\tilde{X}_1 \cap \tilde{X}_2))
$$

=
$$
\mu(\tilde{X}_1 \cap \tilde{X}_2) + \mu(\tilde{X}_1 \cap \tilde{X}_2^c)
$$

