Evaluation and comparison of set-valued estimators: empirical and structural aspects

Evaluation vs Properties

Evaluation

Evaluation vs Properties

Evaluation

in an empirical sense:

Evaluation vs Properties

Evaluation

in an empirical sense:

is able to solve the problem at hand:

Evaluation vs Properties

Evaluation

in an empirical sense:

is able to solve the problem at hand:

ideally has to make true and (fully)

informative statements/predictions:

Evaluation vs Properties

Evaluation
in an empirical sense:
is able to solve the problem at hand:
ideally has to make true and (fully)
informative statements/predictions:
has to be nearly as useful as truth.

Evaluation vs Properties

Evaluation

in an empirical sense:

is able to solve the problem at hand:

ideally has to make true and (fully)

informative statements/predictions:

has to be nearly as useful as truth.

Properties

Evaluation vs Properties

Evaluation

in an empirical sense:

is able to solve the problem at hand:

ideally has to make true and (fully)

informative statements/predictions:

has to be nearly as useful as truth.

Properties

in a structural sense:

Evaluation vs Properties

Evaluation

in an empirical sense:

is able to solve the problem at hand:

ideally has to make true and (fully)

informative statements/predictions:

has to be nearly as useful as truth.

Properties

in a structural sense:

has to have a similar structure like

truth.

Evaluation

Evaluation

Trueness

Evaluation

Trueness

„nearly“ true:

Evaluation

Trueness

„nearly“ true:

- *true only with a certain probability*

Evaluation

Trueness

„nearly“ true:

- *true only with a certain probability*
- *true in a fuzzy sense*

Evaluation

Trueness

„nearly“ true:

- *true only with a certain probability*
- *true in a fuzzy sense*

informativeness:

Evaluation

Trueness

„nearly" true:

- *true only with a certain probability*
- *true in a fuzzy sense*

informativeness: Statements/predictions have to be non-tautological.

How to formalize?

How to formalize?

different approaches:

How to formalize?

different approaches:

- *fixed informativeness:*

How to formalize?

different approaches:

- *fixed informativeness: find estimator with fixed informativeness and best trueness compared to other estimators with the same informativeness*

How to formalize?

different approaches:

- *fixed informativeness: find estimator with fixed informativeness and best trueness compared to other estimators with the same informativeness (e.g. maximal informativeness*

How to formalize?

different approaches:

- *fixed informativeness: find estimator with fixed informativeness and best trueness compared to other estimators with the same informativeness (e.g. maximal informativeness $\Longrightarrow$ point-estimator).*

How to formalize?

different approaches:

- *fixed informativeness: find estimator with fixed informativeness and best trueness compared to other estimators with the same informativeness (e.g. maximal informativeness $\Longrightarrow$ point-estimator).*
- *fixed trueness (e.g. covering probability):*

How to formalize?
different approaches:

- *fixed informativeness: find estimator with fixed informativeness and best trueness compared to other estimators with the same informativeness (e.g. maximal informativeness $\Longrightarrow$ point-estimator).*
- *fixed trueness (e.g. covering probability): find estimator with fixed covering-probability and best informativeness compared to other estimators with the same covering probability.*

How to formalize?
different approaches:

- *fixed informativeness: find estimator with fixed informativeness and best trueness compared to other estimators with the same informativeness (e.g. maximal informativeness $\Longrightarrow$ point-estimator).*
- *fixed trueness (e.g. covering probability): find estimator with fixed covering-probability and best informativeness compared to other estimators with the same covering probability.*
- *„mixtures“ of the first and the second approach?*

Often there is no best estimator, only undominated estimators

second approach:

second approach:

how to measure informativeness?

second approach:

how to measure informativeness?

example:

second approach:

how to measure informativeness?

example: simple linear regression:

second approach:

how to measure informativeness?

example: simple linear regression:

a (set-valued) estimate (e.g. a confidence ellipse) s_1 *is more informative than another estimate* s_2

second approach:

how to measure informativeness?

example: simple linear regression:

a (set-valued) estimate (e.g. a confidence ellipse) s_1 *is more informative than another estimate* s_2 *if* $s_1 \subseteq s_2$.

second approach:

how to measure informativeness?

example: simple linear regression:

a (set-valued) estimate (e.g. a confidence ellipse) s_1 *is more informative than another estimate* s_2 *if* $s_1 \subseteq s_2$.

or accordingly an estimator S_1 *is more informative than* S_2 *if* $S_1(x) \subseteq S_2(x)$ *for all data* x.

second approach:

how to measure informativeness?

example: simple linear regression:

a (set-valued) estimate (e.g. a confidence ellipse) s_1 *is more informative than another estimate* s_2 *if* $s_1 \subseteq s_2$.

or accordingly an estimator S_1 *is more informative than* S_2 *if* $S_1(x) \subseteq S_2(x)$ *for all data* x. *Often:* $s_1 \not\subseteq s_2$ *and* $s_2 \not\subseteq s_1$.

second approach:

how to measure informativeness?

example: simple linear regression:

a (set-valued) estimate (e.g. a confidence ellipse) s_1 is more informative than another estimate s_2 if $s_1 \subseteq s_2$.

or accordingly an estimator S_1 is more informative than S_2 if $S_1(x) \subseteq S_2(x)$ for all data x. Often: $s_1 \nsubseteq s_2$ and $s_2 \nsubseteq s_1$.

Which estimate is more informative?

second approach:

how to measure informativeness?

example: simple linear regression:

a (set-valued) estimate (e.g. a confidence ellipse) s_1 *is more informative than another estimate* s_2 *if* $s_1 \subseteq s_2$.

or accordingly an estimator S_1 *is more informative than* S_2 *if* $S_1(x) \subseteq S_2(x)$ *for all data* x. *Often:* $s_1 \nsubseteq s_2$ *and* $s_2 \nsubseteq s_1$.

Which estimate is more informative?

maybe no one, but in practice we are forced to choose one estimator.

second approach:

how to measure informativeness?

example: simple linear regression:

a (set-valued) estimate (e.g. a confidence ellipse) s_1 is more informative than another estimate s_2 if $s_1 \subseteq s_2$.

or accordingly an estimator S_1 is more informative than S_2 if $S_1(x) \subseteq S_2(x)$ for all data x. Often: $s_1 \nsubseteq s_2$ and $s_2 \nsubseteq s_1$.

Which estimate is more informative?

maybe no one, but in practice we are forced to choose one estimator.

one possible (naive) answer:

second approach:

how to measure informativeness?

example: simple linear regression:

a (set-valued) estimate (e.g. a confidence ellipse) s_1 *is more informative than another estimate* s_2 *if* $s_1 \subseteq s_2$.

or accordingly an estimator S_1 *is more informative than* S_2 *if* $S_1(x) \subseteq S_2(x)$ *for all data* x*. Often:* $s_1 \nsubseteq s_2$ *and* $s_2 \nsubseteq s_1$.

Which estimate is more informative?

maybe no one, but in practice we are forced to choose one estimator.

one possible (naive) answer: the estimate with the lowest area (if this exists)

second approach:

how to measure informativeness?

example: simple linear regression:

a (set-valued) estimate (e.g. a confidence ellipse) s_1 *is more informative than another estimate* s_2 *if* $s_1 \subseteq s_2$.

or accordingly an estimator S_1 *is more informative than* S_2 *if* $S_1(x) \subseteq S_2(x)$ *for all data* x. *Often:* $s_1 \nsubseteq s_2$ *and* $s_2 \nsubseteq s_1$.

Which estimate is more informative?

maybe no one, but in practice we are forced to choose one estimator.

one possible (naive) answer: the estimate with the lowest area (if this exists) or the estimator with the lower expected area (independent from the true parameter) respectively.

Problem with the area

Problem with the area

- *no good interpretation (especially with more than one covariate)*

Problem with the area

- *no good interpretation (especially with more than one covariate)*
- *not invariant under nonlinear transformations of the covariates*

Another way of measuring the informativeness:

Another way of measuring the informativeness:

look on the usefulness of the estimator:

Another way of measuring the informativeness:

look on the usefulness of the estimator: especially the predictive accuracy.

Another way of measuring the informativeness:

look on the usefulness of the estimator: especially the predictive accuracy.

- *first problem:*

Another way of measuring the informativeness:

look on the usefulness of the estimator: especially the predictive accuracy.

- *first problem: loss-function*

Another way of measuring the informativeness:

look on the usefulness of the estimator: especially the predictive accuracy.

- *first problem: loss-function*
- *second problem:*

Another way of measuring the informativeness:
look on the usefulness of the estimator: especially the predictive accuracy.

- *first problem: loss-function*
- *second problem: for which covariates the prediction is evaluated?*

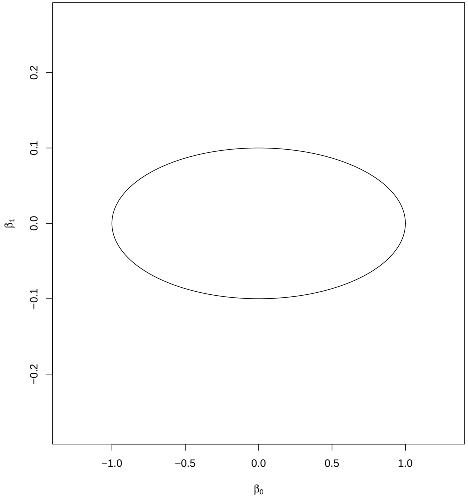
- *also not invariant under nonlinear transformations of the*

- *also not invariant under nonlinear transformations of the dependend variable*

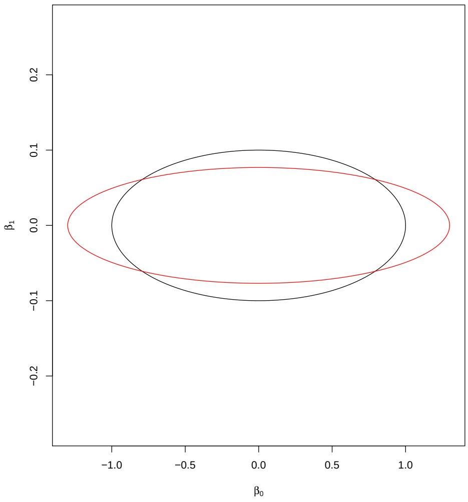
- *also not invariant under nonlinear transformations of the dependend variable (this is entailed in the choice of the loss function)*

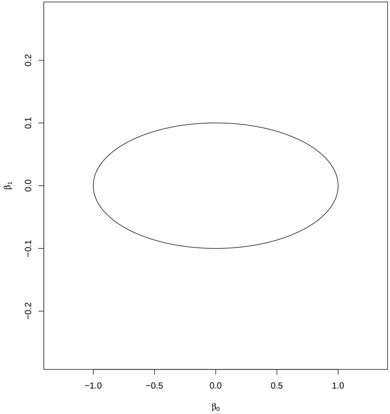
- *also not invariant under nonlinear transformations of the dependend variable (this is entailed in the choice of the loss function)*
- *strange example:*

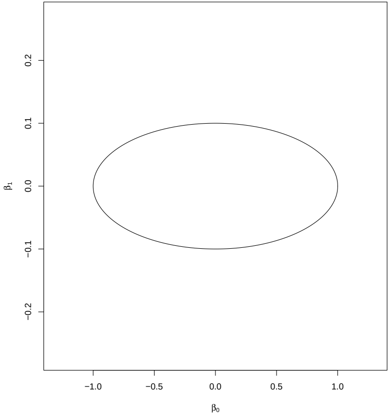
- *also not invariant under nonlinear transformations of the dependend variable (this is entailed in the choice of the loss function)*
- *strange example:*

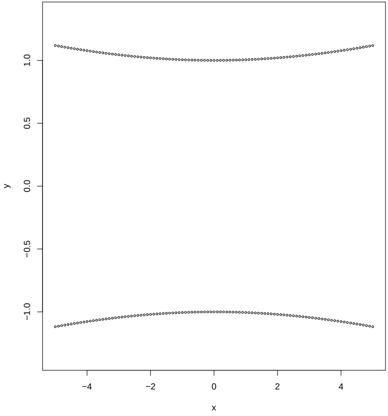


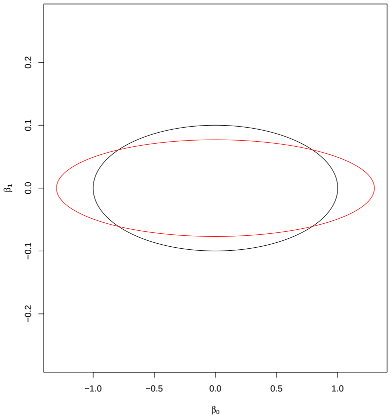
0.2
0.1
0.0
−0.1
−0.2
β_1
−1.0
−0.5
0.0
0.5
1.0
β_0

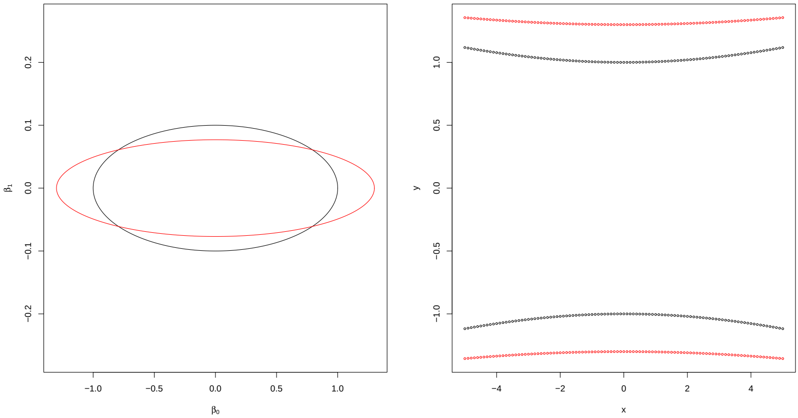

0.2
0.1
0.0
−0.1
−0.2
β1
−1.0
−0.5
0.0
0.5
1.0
β0


0.2
0.1
0.0
−0.1
−0.2
β_1
−1.0
−0.5
0.0
0.5
1.0
β_0

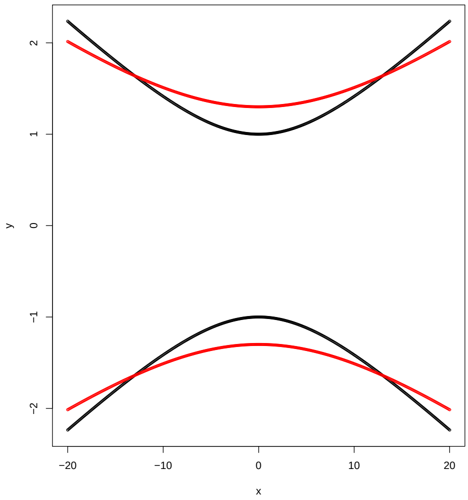

β1
0.2
0.1
0.0
−0.1
−0.2
−1.0
−0.5
0.0
0.5
1.0
β0

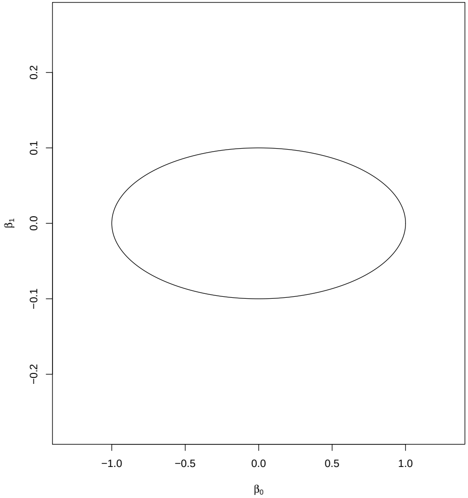

y
1.0
0.5
0.0
−0.5
−1.0
−4
−2
0
2
4
x


β1
0.2
0.1
0.0
−0.1
−0.2
−1.0
−0.5
0.0
0.5
1.0
β0
y
1.0
0.5
0.0
−0.5
−1.0
−4
−2
0
2
4
x

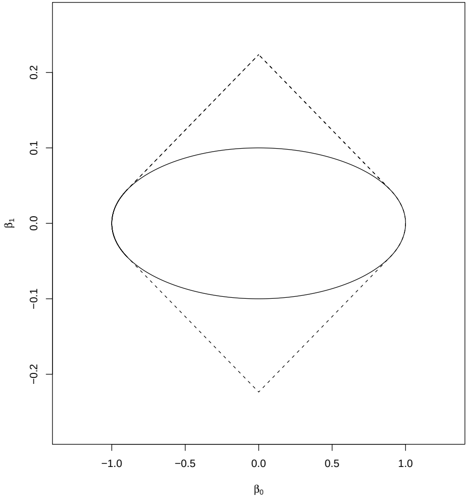


0.2
0.1
0.0
−0.1
−0.2
β_1
−1.0
−0.5
0.0
0.5
1.0
β_0
1.0
0.5
0.0
−0.5
−1.0
y
−4
−2
0
2
4
x

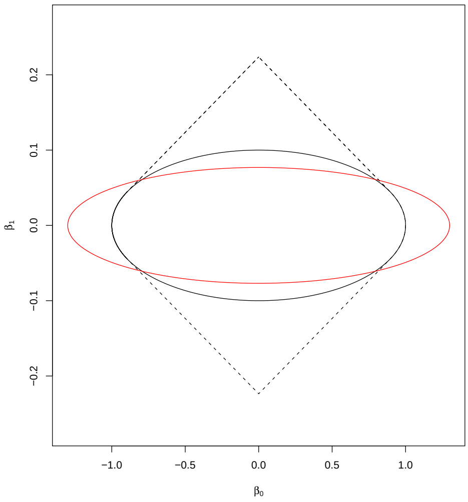

y
2
1
0
−1
−2
−20
−10
0
10
20
x

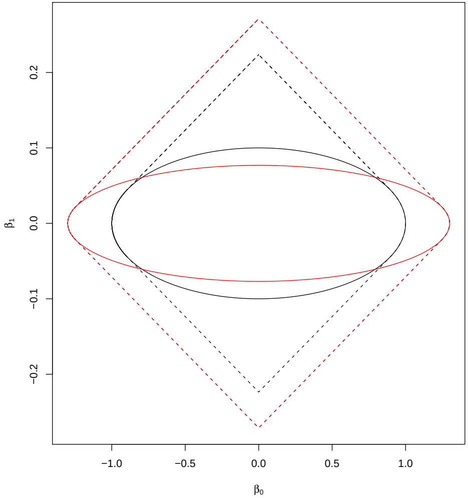


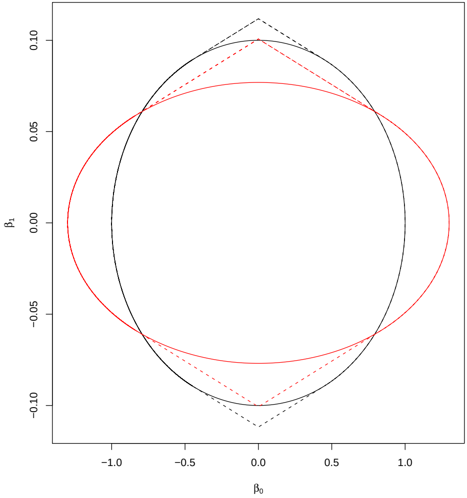
0.2
0.1
0.0
−0.1
−0.2
β_1
−1.0
−0.5
0.0
0.5
1.0
β_0



0.2
0.1
0.0
−0.1
−0.2
β_1
−1.0
−0.5
0.0
0.5
1.0
β_0


0.2
0.1
0.0
−0.1
−0.2
β_1
−1.0
−0.5
0.0
0.5
1.0
β_0


0.2
0.1
0.0
−0.1
−0.2
β1
−1.0
−0.5
0.0
0.5
1.0
β0


0.10
0.05
0.00
−0.05
−0.10
β_1
−1.0
−0.5
0.0
0.5
1.0
β_0

Structural Properties

Structural Properties

idea:

Structural Properties

idea: an estimator, as an „approximation of the truth" should have a „similar structure like the truth". simple example:

Structural Properties

idea: an estimator, as an „approximation of the truth" should have a „similar structure like the truth". simple example:location estimator (point-estimator):

Structural Properties

idea: an estimator, as an „approximation of the truth" should have a „similar structure like the truth". simple example:location estimator (point-estimator):

$e : \mathbb{R}^n :\longrightarrow \mathbb{R}$

Structural Properties

idea: an estimator, as an „approximation of the truth" should have a „similar structure like the truth". simple example:location estimator (point-estimator):

$e : \mathbb{R}^n :\longrightarrow \mathbb{R}$ *, or as a predictor:*

Structural Properties

idea: an estimator, as an „approximation of the truth" should have a „similar structure like the truth". simple example:location estimator (point-estimator):

$e : \mathbb{R}^n :\longrightarrow \mathbb{R}$ *, or as a predictor:*

$e : \mathbb{R}^n \longrightarrow \mathbb{R}^n$.

The mean, as an estimator of the expectation of a random variable.

The mean, as an estimator of the expectation of a random variable.

The mean, as an estimator of the expectation of a random variable.

properties of the mean, as well as the expectation:

The mean, as an estimator of the expectation of a random variable.

properties of the mean, as well as the expectation:

a) *linear:* $e(a \cdot x + y) = a \cdot e(x) + e(y)$

The mean, as an estimator of the expectation of a random variable.

properties of the mean, as well as the expectation:

a) *linear:* $e(a \cdot x + y) = a \cdot e(x) + e(y)$

a') *especially:*

The mean, as an estimator of the expectation of a random variable.

properties of the mean, as well as the expectation:

a) *linear:* $e(a \cdot x + y) = a \cdot e(x) + e(y)$

a') *especially: affine equivariant:* $e(a \cdot x + b) = a \cdot e(x) + b$

The mean, as an estimator of the expectation of a random variable.

properties of the mean, as well as the expectation:

a) *linear:* $e(a \cdot x + y) = a \cdot e(x) + e(y)$

a') *especially: affine equivariant:* $e(a \cdot x + b) = a \cdot e(x) + b$

b) *monotone:*

The mean, as an estimator of the expectation of a random variable.

properties of the mean, as well as the expectation:

a) *linear:* $e(a \cdot x + y) = a \cdot e(x) + e(y)$

a') *especially: affine equivariant:* $e(a \cdot x + b) = a \cdot e(x) + b$

b) *monotone:* $x \leq y \Longrightarrow e(x) \leq e(y)$

The mean, as an estimator of the expectation of a random variable.

properties of the mean, as well as the expectation:

a) *linear:* $e(a \cdot x + y) = a \cdot e(x) + e(y)$

a') *especially: affine equivariant:* $e(a \cdot x + b) = a \cdot e(x) + b$

b) *monotone:* $x \leq y \Longrightarrow e(x) \leq e(y)$

easy to compute image of Intervals:

The mean, as an estimator of the expectation of a random variable.

properties of the mean, as well as the expectation:

a) *linear:* $e(a \cdot x + y) = a \cdot e(x) + e(y)$

a') *especially: affine equivariant:* $e(a \cdot x + b) = a \cdot e(x) + b$

b) *monotone:* $x \leq y \Longrightarrow e(x) \leq e(y)$
easy to compute image of Intervals:
$I = \{y | \underline{y} \leq y \leq \overline{y}\}$

The mean, as an estimator of the expectation of a random variable.

properties of the mean, as well as the expectation:

a) *linear:* $e(a \cdot x + y) = a \cdot e(x) + e(y)$

a') *especially: affine equivariant:* $e(a \cdot x + b) = a \cdot e(x) + b$

b) *monotone:* $x \leq y \Longrightarrow e(x) \leq e(y)$

easy to compute image of Intervals:

$I = \{y | \underline{y} \leq y \leq \overline{y}\} \Longrightarrow e[I] = \{z | e(\underline{y}) \leq z \leq e(\overline{y})\}$

The mean, as an estimator of the expectation of a random variable.

properties of the mean, as well as the expectation:

a) *linear:* $e(a \cdot x + y) = a \cdot e(x) + e(y)$

a') *especially: affine equivariant:* $e(a \cdot x + b) = a \cdot e(x) + b$

b) *monotone:* $x \leq y \Longrightarrow e(x) \leq e(y)$

easy to compute image of Intervals:

$I = \{y | \underline{y} \leq y \leq \overline{y}\} \Longrightarrow e[I] = \{z | e(\underline{y}) \leq z \leq e(\overline{y})\}$

c) *idempotent:*

The mean, as an estimator of the expectation of a random variable.

properties of the mean, as well as the expectation:

a) *linear:* $e(a \cdot x + y) = a \cdot e(x) + e(y)$

a') *especially: affine equivariant:* $e(a \cdot x + b) = a \cdot e(x) + b$

b) *monotone:* $x \leq y \Longrightarrow e(x) \leq e(y)$

easy to compute image of Intervals:

$I = \{y | \underline{y} \leq y \leq \overline{y}\} \Longrightarrow e[I] = \{z | e(\underline{y}) \leq z \leq e(\overline{y})\}$

c) *idempotent:* $e(e(x)) = e(x)$

The mean, as an estimator of the expectation of a random variable.

properties of the mean, as well as the expectation:

a) *linear:* $e(a \cdot x + y) = a \cdot e(x) + e(y)$

a') *especially: affine equivariant:* $e(a \cdot x + b) = a \cdot e(x) + b$

b) *monotone:* $x \leq y \Longrightarrow e(x) \leq e(y)$

easy to compute image of Intervals:

$I = \{y | \underline{y} \leq y \leq \overline{y}\} \Longrightarrow e[I] = \{z | e(\underline{y}) \leq z \leq e(\overline{y})\}$

c) *idempotent:* $e(e(x)) = e(x)$

d) *anti-extensive:* $e(x) \geq x$

The mean, as an estimator of the expectation of a random variable.

properties of the mean, as well as the expectation:

a) *linear:* $e(a \cdot x + y) = a \cdot e(x) + e(y)$

a') *especially: affine equivariant:* $e(a \cdot x + b) = a \cdot e(x) + b$

b) *monotone:* $x \leq y \Longrightarrow e(x) \leq e(y)$

easy to compute image of Intervals:

$I = \{y | \underline{y} \leq y \leq \overline{y}\} \Longrightarrow e[I] = \{z | e(\underline{y}) \leq z \leq e(\overline{y})\}$

c) *idempotent:* $e(e(x)) = e(x)$

d) *anti-extensive:* $e(x) \geq x \Longrightarrow e(x) = x$

The mean, as an estimator of the expectation of a random variable.

properties of the mean, as well as the expectation:

a) *linear:* $e(a \cdot x + y) = a \cdot e(x) + e(y)$

a') *especially: affine equivariant:* $e(a \cdot x + b) = a \cdot e(x) + b$

b) *monotone:* $x \leq y \Longrightarrow e(x) \leq e(y)$
easy to compute image of Intervals:
$I = \{y | \underline{y} \leq y \leq \overline{y}\} \Longrightarrow e[I] = \{z | e(\underline{y}) \leq z \leq e(\overline{y})\}$

c) *idempotent:* $e(e(x)) = e(x)$

d) *anti-extensive:* $e(x) \geq x \Longrightarrow e(x) = x$

d') *anti-intensive:* $e(x) \leq x$

The mean, as an estimator of the expectation of a random variable.

properties of the mean, as well as the expectation:

a) *linear:* $e(a \cdot x + y) = a \cdot e(x) + e(y)$

a') *especially: affine equivariant:* $e(a \cdot x + b) = a \cdot e(x) + b$

b) *monotone:* $x \leq y \Longrightarrow e(x) \leq e(y)$

easy to compute image of Intervals:

$I = \{y | \underline{y} \leq y \leq \overline{y}\} \Longrightarrow e[I] = \{z | e(\underline{y}) \leq z \leq e(\overline{y})\}$

c) *idempotent:* $e(e(x)) = e(x)$

d) *anti-extensive:* $e(x) \geq x \Longrightarrow e(x) = x$

d') *anti-intensive:* $e(x) \leq x \Longrightarrow e(x) = x$

Most location parameters (e.g. median, linearly weighted mean, winsorized mean, Hodges-Lehmann-estimator) also meet the properties a') till d).

set-valued estimators

set-valued estimators

Generalization to set-valued location estimators:

set-valued estimators

Generalization to set-valued location estimators:

$e : \mathbb{R}^n \longrightarrow 2^{\mathbb{R}^n}$

set-valued estimators

Generalization to set-valued location estimators:

$e : \mathbb{R}^n \longrightarrow 2^{\mathbb{R}^n}$ *(, or especially interval-valued estimators).*

set-valued estimators

Generalization to set-valued location estimators:

$e : \mathbb{R}^n \longrightarrow 2^{\mathbb{R}^n}$ *(, or especially interval-valued estimators).*

a') *affine equivariant:* $e(a \cdot x + b) = a \cdot e(x) + b$

set-valued estimators

Generalization to set-valued location estimators:
$e : \mathbb{R}^n \longrightarrow 2^{\mathbb{R}^n}$ *(, or especially interval-valued estimators).*

a') *affine equivariant:* $e(a \cdot x + b) = a \cdot e(x) + b$

b) *monotone:*

set-valued estimators

Generalization to set-valued location estimators:

$e : \mathbb{R}^n \longrightarrow 2^{\mathbb{R}^n}$ *(, or especially interval-valued estimators).*

a') *affine equivariant:* $e(a \cdot x + b) = a \cdot e(x) + b$

b) *monotone:* $x \leq y \Longrightarrow \underline{e}(x) \leq \underline{e}(y)$ & $\overline{e}(x) \leq \overline{e}(y)$

set-valued estimators

Generalization to set-valued location estimators:

$e : \mathbb{R}^n \longrightarrow 2^{\mathbb{R}^n}$ *(, or especially interval-valued estimators).*

a') *affine equivariant:* $e(a \cdot x + b) = a \cdot e(x) + b$

b) *monotone:* $x \leq y \Longrightarrow \underline{e}(x) \leq \underline{e}(y) \;\;\&\;\; \overline{e}(x) \leq \overline{e}(y)$

easy to compute image of Intervals:

set-valued estimators

Generalization to set-valued location estimators:
$e : \mathbb{R}^n \longrightarrow 2^{\mathbb{R}^n}$ *(, or especially interval-valued estimators).*

a') *affine equivariant:* $e(a \cdot x + b) = a \cdot e(x) + b$

b) *monotone:* $x \leq y \Longrightarrow \underline{\mathrm{e}}(x) \leq \underline{\mathrm{e}}(y)$ & $\overline{\mathrm{e}}(x) \leq \overline{\mathrm{e}}(y)$
easy to compute image of Intervals:
$I = \{y | \underline{\mathrm{y}} \leq y \leq \overline{\mathrm{y}}\}$

set-valued estimators

Generalization to set-valued location estimators:
$e : \mathbb{R}^n \longrightarrow 2^{\mathbb{R}^n}$ *(, or especially interval-valued estimators).*

a') *affine equivariant:* $e(a \cdot x + b) = a \cdot e(x) + b$

b) *monotone:* $x \leq y \Longrightarrow \underline{e}(x) \leq \underline{e}(y)$ & $\overline{e}(x) \leq \overline{e}(y)$
easy to compute image of Intervals:
$I = \{y \mid \underline{y} \leq y \leq \overline{y}\} \Longrightarrow e[I] = \{z \mid \underline{e}(\underline{y}) \leq z \leq \overline{e}(\overline{y})\}$

set-valued estimators

Generalization to set-valued location estimators:
$e : \mathbb{R}^n \longrightarrow 2^{\mathbb{R}^n}$ *(, or especially interval-valued estimators).*

a') *affine equivariant:* $e(a \cdot x + b) = a \cdot e(x) + b$

b) *monotone:* $x \leq y \Longrightarrow \underline{e}(x) \leq \underline{e}(y)$ & $\overline{e}(x) \leq \overline{e}(y)$
easy to compute image of Intervals:
$I = \{y | \underline{y} \leq y \leq \overline{y}\} \Longrightarrow e[I] = \{z | \underline{e}(\underline{y}) \leq z \leq \overline{e}(\overline{y})\}$

set-valued estimators

Generalization to set-valued location estimators:
$e : \mathbb{R}^n \longrightarrow 2^{\mathbb{R}^n}$ *(, or especially interval-valued estimators).*

a') *affine equivariant:* $e(a \cdot x + b) = a \cdot e(x) + b$

b) *monotone:* $x \leq y \Longrightarrow \underline{e}(x) \leq \underline{e}(y)$ & $\overline{e}(x) \leq \overline{e}(y)$
easy to compute image of Intervals:
$I = \{y | \underline{y} \leq y \leq \overline{y}\} \Longrightarrow e[I] = \{z | \underline{e}(\underline{y}) \leq z \leq \overline{e}(\overline{y})\}$

d) *anti-extensive:* $\underline{e}(x) \geq x$

set-valued estimators

Generalization to set-valued location estimators:
$e : \mathbb{R}^n \longrightarrow 2^{\mathbb{R}^n}$ *(, or especially interval-valued estimators).*

a') *affine equivariant:* $e(a \cdot x + b) = a \cdot e(x) + b$

b) *monotone:* $x \leq y \Longrightarrow \underline{e}(x) \leq \underline{e}(y)$ & $\overline{e}(x) \leq \overline{e}(y)$
easy to compute image of Intervals:
$I = \{y | \underline{y} \leq y \leq \overline{y}\} \Longrightarrow e[I] = \{z | \underline{e}(\underline{y}) \leq z \leq \overline{e}(\overline{y})\}$

d) *anti-extensive:* $\underline{e}(x) \geq x \Longrightarrow \underline{e}(x) = \overline{e}(x) = x$

set-valued estimators

Generalization to set-valued location estimators:
$e : \mathbb{R}^n \longrightarrow 2^{\mathbb{R}^n}$ *(, or especially interval-valued estimators).*

a') *affine equivariant:* $e(a \cdot x + b) = a \cdot e(x) + b$

b) *monotone:* $x \leq y \Longrightarrow \underline{e}(x) \leq \underline{e}(y)$ & $\overline{e}(x) \leq \overline{e}(y)$
easy to compute image of Intervals:
$I = \{y | \underline{y} \leq y \leq \overline{y}\} \Longrightarrow e[I] = \{z | \underline{e}(\underline{y}) \leq z \leq \overline{e}(\overline{y})\}$

d) *anti-extensive:* $\underline{e}(x) \geq x \Longrightarrow \underline{e}(x) = \overline{e}(x) = x$

d') *anti-intensive:* $\overline{e}(x) \leq x$

set-valued estimators

Generalization to set-valued location estimators:
$e : \mathbb{R}^n \longrightarrow 2^{\mathbb{R}^n}$ *(, or especially interval-valued estimators).*

a') *affine equivariant:* $e(a \cdot x + b) = a \cdot e(x) + b$

b) *monotone:* $x \leq y \Longrightarrow \underline{e}(x) \leq \underline{e}(y)$ & $\overline{e}(x) \leq \overline{e}(y)$
easy to compute image of Intervals:
$I = \{y \mid \underline{y} \leq y \leq \overline{y}\} \Longrightarrow e[I] = \{z \mid \underline{e}(\underline{y}) \leq z \leq \overline{e}(\overline{y})\}$

d) *anti-extensive:* $\underline{e}(x) \geq x \Longrightarrow \underline{e}(x) = \overline{e}(x) = x$

d') *anti-intensive:* $\overline{e}(x) \leq x \Longrightarrow \underline{e}(x) = \overline{e}(x) = x$

set-valued estimators

Generalization to set-valued location estimators:
$e : \mathbb{R}^n \longrightarrow 2^{\mathbb{R}^n}$ *(, or especially interval-valued estimators).*

a') *affine equivariant:* $e(a \cdot x + b) = a \cdot e(x) + b$

b) *monotone:* $x \leq y \Longrightarrow \underline{e}(x) \leq \underline{e}(y)$ & $\overline{e}(x) \leq \overline{e}(y)$
easy to compute image of Intervals:
$I = \{y | \underline{y} \leq y \leq \overline{y}\} \Longrightarrow e[I] = \{z | \underline{e}(\underline{y}) \leq z \leq \overline{e}(\overline{y})\}$

d) *anti-extensive:* $\underline{e}(x) \geq x \Longrightarrow \underline{e}(x) = \overline{e}(x) = x$

d') *anti-intensive:* $\overline{e}(x) \leq x \Longrightarrow \underline{e}(x) = \overline{e}(x) = x$

c) *idempotence:*

set-valued estimators

Generalization to set-valued location estimators:
$e : \mathbb{R}^n \longrightarrow 2^{\mathbb{R}^n}$ *(, or especially interval-valued estimators).*

a') *affine equivariant:* $e(a \cdot x + b) = a \cdot e(x) + b$

b) *monotone:* $x \leq y \Longrightarrow \underline{e}(x) \leq \underline{e}(y)$ & $\overline{e}(x) \leq \overline{e}(y)$
easy to compute image of Intervals:
$I = \{y | \underline{y} \leq y \leq \overline{y}\} \Longrightarrow e[I] = \{z | \underline{e}(\underline{y}) \leq z \leq \overline{e}(\overline{y})\}$

d) *anti-extensive:* $\underline{e}(x) \geq x \Longrightarrow \underline{e}(x) = \overline{e}(x) = x$

d') *anti-intensive:* $\overline{e}(x) \leq x \Longrightarrow \underline{e}(x) = \overline{e}(x) = x$

c) *idempotence: makes more sense for set domained, set valued estimators.*

set-domained and set-valued estimators

set-domained and set-valued estimators

Generalization to set-domained, set-valued location estimators

set-domained and set-valued estimators

Generalization to set-domained, set-valued location estimators
$e : 2^{\mathbb{R}^n} \longrightarrow 2^{\mathbb{R}^n}$

set-domained and set-valued estimators

Generalization to set-domained, set-valued location estimators $e : 2^{\mathbb{R}^n} \longrightarrow 2^{\mathbb{R}^n}$ *(, or especially interval-domained, interval-valued estimators):*

set-domained and set-valued estimators

Generalization to set-domained, set-valued location estimators $e : 2^{\mathbb{R}^n} \longrightarrow 2^{\mathbb{R}^n}$ (, or especially interval-domained, interval-valued estimators):

c) *idempotence:*

c) *idempotence:* $e(e(X)) = X$

a') *affine equivariant:* $e(a \cdot X + b) = a \cdot e(X) + b$

c) *idempotence:* $e(e(X)) = X$

a') *affine equivariant:* $e(a \cdot X + b) = a \cdot e(X) + b$

b) *(elementwise) monotone:*

c) *idempotence:* $e(e(X)) = X$

a') *affine equivariant:* $e(a \cdot X + b) = a \cdot e(X) + b$

b) *(elementwise) monotone:*
$x \leq y \Longrightarrow \underline{e}(\{x\}) \leq \underline{e}(\{y\}) \;\; \& \;\; \overline{e}(\{x\}) \leq \overline{e}(\{y\})$

c) *idempotence:* $e(e(X)) = X$

a') *affine equivariant:* $e(a \cdot X + b) = a \cdot e(X) + b$

b) *(elementwise) monotone:*
$x \leq y \Longrightarrow \underline{e}(\{x\}) \leq \underline{e}(\{y\}) \;\; \& \;\; \overline{e}(\{x\}) \leq \overline{e}(\{y\})$

e) *set-monotone:*

c) *idempotence:* $e(e(X)) = X$

a') *affine equivariant:* $e(a \cdot X + b) = a \cdot e(X) + b$

b) *(elementwise) monotone:*
$x \leq y \Longrightarrow \underline{e}(\{x\}) \leq \underline{e}(\{y\})\ \ \&\ \ \overline{e}(\{x\}) \leq \overline{e}(\{y\})$

e) *set-monotone:* $X \subseteq Y$

c) *idempotence:* $e(e(X)) = X$

a') *affine equivariant:* $e(a \cdot X + b) = a \cdot e(X) + b$

b) *(elementwise) monotone:*
$x \leq y \Longrightarrow \underline{e}(\{x\}) \leq \underline{e}(\{y\}) \;\; \& \;\; \overline{e}(\{x\}) \leq \overline{e}(\{y\})$

e) *set-monotone:* $X \subseteq Y \Longrightarrow e(X) \subseteq e(Y)$

c) *idempotence:* $e(e(X)) = X$

a') *affine equivariant:* $e(a \cdot X + b) = a \cdot e(X) + b$

b) *(elementwise) monotone:*
$x \leq y \Longrightarrow \underline{e}(\{x\}) \leq \underline{e}(\{y\}) \;\;\&\;\; \overline{e}(\{x\}) \leq \overline{e}(\{y\})$

e) *set-monotone:* $X \subseteq Y \Longrightarrow e(X) \subseteq e(Y)$
for all $x \in X$ *it follows:*

c) *idempotence:* $e(e(X)) = X$

a') *affine equivariant:* $e(a \cdot X + b) = a \cdot e(X) + b$

b) *(elementwise) monotone:*
$x \leq y \Longrightarrow \underline{e}(\{x\}) \leq \underline{e}(\{y\}) \;\; \& \;\; \overline{e}(\{x\}) \leq \overline{e}(\{y\})$

e) *set-monotone:* $X \subseteq Y \Longrightarrow e(X) \subseteq e(Y)$
for all $x \in X$ *it follows:*

$$\{x\} \quad \subseteq \bigcup_{x \in X} \{x\} = X$$

c) *idempotence:* $e(e(X)) = X$

a') *affine equivariant:* $e(a \cdot X + b) = a \cdot e(X) + b$

b) *(elementwise) monotone:*
$x \leq y \Longrightarrow \underline{e}(\{x\}) \leq \underline{e}(\{y\}) \;\; \& \;\; \overline{e}(\{x\}) \leq \overline{e}(\{y\})$

e) *set-monotone:* $X \subseteq Y \Longrightarrow e(X) \subseteq e(Y)$
for all $x \in X$ *it follows:*

$$\begin{aligned} \{x\} &\subseteq \bigcup_{x \in X} \{x\} = X \\ \Longrightarrow &\quad e(\{x\}) \subseteq e(X) \end{aligned}$$

c) *idempotence:* $e(e(X)) = X$

a') *affine equivariant:* $e(a \cdot X + b) = a \cdot e(X) + b$

b) *(elementwise) monotone:*
$x \leq y \Longrightarrow \underline{e}(\{x\}) \leq \underline{e}(\{y\}) \;\;\&\;\; \overline{e}(\{x\}) \leq \overline{e}(\{y\})$

e) *set-monotone:* $X \subseteq Y \Longrightarrow e(X) \subseteq e(Y)$
for all $x \in X$ *it follows:*

$$\begin{aligned} \{x\} &\subseteq \bigcup_{x \in X} \{x\} = X \\ \Longrightarrow &\quad e(\{x\}) \subseteq e(X) \\ \Longrightarrow &\quad \bigcup_{x \in X} e(\{x\}) \subseteq e(X). \end{aligned}$$

c) *idempotence:* $e(e(X)) = X$

a') *affine equivariant:* $e(a \cdot X + b) = a \cdot e(X) + b$

b) *(elementwise) monotone:*
$x \leq y \Longrightarrow \underline{e}(\{x\}) \leq \underline{e}(\{y\}) \;\; \& \;\; \overline{e}(\{x\}) \leq \overline{e}(\{y\})$

e) *set-monotone:* $X \subseteq Y \Longrightarrow e(X) \subseteq e(Y)$
for all $x \in X$ *it follows:*

$$\begin{aligned} \{x\} &\subseteq \bigcup_{x \in X} \{x\} = X \\ \Longrightarrow &\quad e(\{x\}) \subseteq e(X) \\ \Longrightarrow &\quad \bigcup_{x \in X} e(\{x\}) \subseteq e(X). \end{aligned}$$

f) *compositional:*

c) *idempotence:* $e(e(X)) = X$

a') *affine equivariant:* $e(a \cdot X + b) = a \cdot e(X) + b$

b) *(elementwise) monotone:*
$x \leq y \Longrightarrow \underline{e}(\{x\}) \leq \underline{e}(\{y\}) \;\; \& \;\; \overline{e}(\{x\}) \leq \overline{e}(\{y\})$

e) *set-monotone:* $X \subseteq Y \Longrightarrow e(X) \subseteq e(Y)$
for all $x \in X$ *it follows:*

$$\begin{aligned} \{x\} &\subseteq \bigcup_{x \in X} \{x\} = X \\ \Longrightarrow &\quad e(\{x\}) \subseteq e(X) \\ \Longrightarrow &\quad \bigcup_{x \in X} e(\{x\}) \subseteq e(X). \end{aligned}$$

f) *compositional:* $\bigcup_{x \in X} e(\{x\}) = e(X)$

Why the need to satisfy structural properties?

Why the need to satisfy structural properties?

not so clear.

Why the need to satisfy structural properties?

not so clear.

- *computations and constructions are often easier:*

Why the need to satisfy structural properties?

not so clear.

- *computations and constructions are often easier:*
  for example from an interval-valued, point-domained <u>*monotone*</u> *location-estimator we can easily construct an interval-domained (interval-valued) location-estimator:*
  $X = [\underline{x}, \overline{x}]$

Why the need to satisfy structural properties?

not so clear.

- *computations and constructions are often easier:*
  for example from an interval-valued, point-domained <u>*monotone*</u> *location-estimator we can easily construct an interval-domained (interval-valued) location-estimator:*
  $X = [\underline{x}, \overline{x}] \Longrightarrow e(X) = [\underline{e}(\underline{x}), \overline{e}(\overline{x})]$.

Why the need to satisfy structural properties?

not so clear.

- *computations and constructions are often easier:*
  for example from an interval-valued, point-domained *<u>monotone</u>* *location-estimator we can easily construct an interval-domained (interval-valued) location-estimator:*
  $X = [\underline{x}, \overline{x}] \Longrightarrow e(X) = [\underline{e}(\underline{x}), \overline{e}(\overline{x})]$.
  not so easy to generalize for example a classical confidence-interval
  $CI = [\hat{\mu} - c \cdot \hat{\sigma}, \hat{\mu} + c \cdot \hat{\sigma}]$.

Why the need to satisfy structural properties?

not so clear.

- *computations and constructions are often easier:*
  for example from an interval-valued, point-domained <u>*monotone*</u> *location-estimator we can easily construct an interval-domained (interval-valued) location-estimator:*
  $X = [\underline{x}, \overline{x}] \Longrightarrow e(X) = [\underline{e}(\underline{x}), \overline{e}(\overline{x})]$.
  not so easy to generalize for example a classical confidence-interval
  $CI = [\hat{\mu} - c \cdot \hat{\sigma}, \hat{\mu} + c \cdot \hat{\sigma}]$.
- *it is easier to carry out simulation studies*

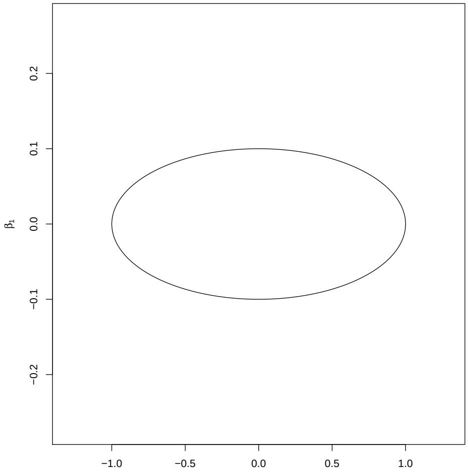
- *as-if-coherence:*

- *as-if-coherence: „estimator should behave, as if it says the truth".*

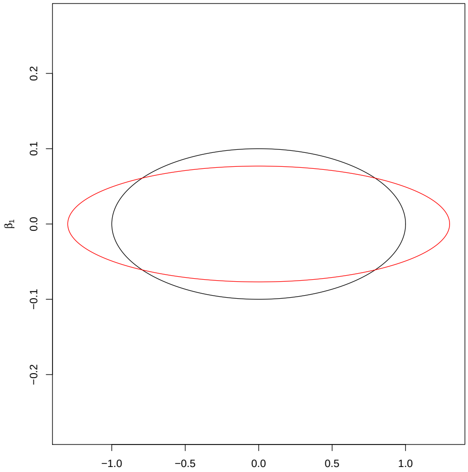
- *as-if-coherence: „estimator should behave, as if it says the truth“.*
  $\Longrightarrow$ *it is maybe not so problematic to deduce propositions from different estimates and ignore the fact, that the estimates are only estimates.*

- *as-if-coherence: „estimator should behave, as if it says the truth“.*
  $\Longrightarrow$ *it is maybe not so problematic to deduce propositions from different estimates and ignore the fact, that the estimates are only estimates.*
  Beispiel?

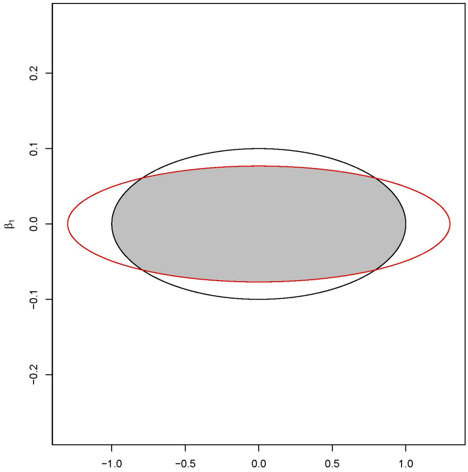
example 1:



example 1:



example 1:



set-monotonicity

set-monotonicity

If we infer sets of possible parameters from sets of possible data,

set-monotonicity

If we infer sets of possible parameters from sets of possible data, from an information-theoretic point of view,

set-monotonicity

If we infer sets of possible parameters from sets of possible data, from an information-theoretic point of view, the following should be intuitively true:

set-monotonicity

If we infer sets of possible parameters from sets of possible data, from an information-theoretic point of view, the following should be intuitively true:
if we have more information about the data (i.e. more precise data, e.g. $X \subseteq Y$), than we should estimate a more informative set of parameters ($e(X) \subseteq e(Y)$).

set-monotonicity

If we infer sets of possible parameters from sets of possible data, from an information-theoretic point of view, the following should be intuitively true:
if we have more information about the data (i.e. more precise data, e.g. $X \subseteq Y$), than we should estimate a more informative set of parameters ($e(X) \subseteq e(Y)$).
What would happen, if this is not the case?

set-monotonicity

If we infer sets of possible parameters from sets of possible data, from an information-theoretic point of view, the following should be intuitively true:
if we have more information about the data (i.e. more precise data, e.g. $X \subseteq Y$), than we should estimate a more informative set of parameters ($e(X) \subseteq e(Y)$).
What would happen, if this is not the case?
if we know, that the true date x lies in X and $X \subseteq Y$, it follows $x \in Y$ and we could „deduce", that the true parameter lies in $e(Y)$, so if $e(X) \not\subseteq e(Y)$ we could sharpen the estimate $e(X)$ to $e^*(X) := e(X) \cap e(Y)$.

set-monotonicity

If we infer sets of possible parameters from sets of possible data, from an information-theoretic point of view, the following should be intuitively true:
if we have more information about the data (i.e. more precise data, e.g. $X \subseteq Y$*), than we should estimate a more informative set of parameters (*$e(X) \subseteq e(Y)$*).*
What would happen, if this is not the case?
if we know, that the true date x *lies in* X *and* $X \subseteq Y$*, it follows* $x \in Y$ *and we could „deduce", that the true parameter lies in* $e(Y)$*, so if* $e(X) \not\subseteq e(Y)$ *we could sharpen the estimate* $e(X)$ *to* $e^*(X) := e(X) \cap e(Y)$*.*
all in all we could construct $e^* : 2^D \longrightarrow 2^P$ *:*

set-monotonicity

If we infer sets of possible parameters from sets of possible data, from an information-theoretic point of view, the following should be intuitively true:
if we have more information about the data (i.e. more precise data, e.g. $X \subseteq Y$), than we should estimate a more informative set of parameters ($e(X) \subseteq e(Y)$).
What would happen, if this is not the case?
if we know, that the true date x lies in X and $X \subseteq Y$, it follows $x \in Y$ and we could „deduce“, that the true parameter lies in $e(Y)$, so if $e(X) \nsubseteq e(Y)$ we could sharpen the estimate $e(X)$ to $e^*(X) := e(X) \cap e(Y)$.
all in all we could construct $e^* : 2^D \longrightarrow 2^P : X \mapsto \bigcap_{Y \supseteq X} e(X)$,

set-monotonicity

If we infer sets of possible parameters from sets of possible data, from an information-theoretic point of view, the following should be intuitively true:
if we have more information about the data (i.e. more precise data, e.g. $X \subseteq Y$), than we should estimate a more informative set of parameters ($e(X) \subseteq e(Y)$).
What would happen, if this is not the case?
if we know, that the true date x lies in X and $X \subseteq Y$, it follows $x \in Y$ and we could „deduce“, that the true parameter lies in $e(Y)$, so if $e(X) \nsubseteq e(Y)$ we could sharpen the estimate $e(X)$ to $e^*(X) := e(X) \cap e(Y)$.
all in all we could construct $e^* : 2^D \longrightarrow 2^P : X \mapsto \bigcap_{Y \supseteq X} e(X)$, which is set-monotone and more informative than e.

compositionality

if e is monotone, but not compositional, one could get, with a similar reasoning, the sharper estimator

compositionality

if e is monotone, but not compositional, one could get, with a similar reasoning, the sharper estimator
$e^{**} : 2^D \longrightarrow 2^P$:

compositionality

if e is monotone, but not compositional, one could get, with a similar reasoning, the sharper estimator

$$e^{**} : 2^D \longrightarrow 2^P : X \mapsto \bigcup_{x \in X} e(\{x\}).$$

Definition

let C be a class of estimators

Definition

let C be a class of estimators (e.g. all estimators $e : 2^D \longrightarrow 2^P$ with a coveringprobability not lower than $1 - \alpha$).

Definition

let C be a class of estimators (e.g. all estimators $e : 2^D \longrightarrow 2^P$ with a coveringprobability not lower than $1 - \alpha$).
An estimator $e \in C$ is called undominated on $K \subseteq 2^D$ with respect to the class C,

Definition

let C be a class of estimators (e.g. all estimators $e : 2^D \longrightarrow 2^P$ with a coveringprobability not lower than $1 - \alpha$).
An estimator $e \in C$ is called undominated on $K \subseteq 2^D$ with respect to the class C,if for all other $e^* \in C \backslash \{e\}$ there exists a set $X \in K$ with

Definition

let C be a class of estimators (e.g. all estimators $e : 2^D \longrightarrow 2^P$ with a coveringprobability not lower than $1 - \alpha$).
An estimator $e \in C$ is called undominated on $K \subseteq 2^D$ with respect to the class C,if for all other $e^* \in C \setminus \{e\}$ there exists a set $X \in K$ with $e^*(X) \nsubseteq e(X)$.

Definition

let C be a class of estimators (e.g. all estimators $e : 2^D \longrightarrow 2^P$ with a coveringprobability not lower than $1-\alpha$).
An estimator $e \in C$ is called undominated on $K \subseteq 2^D$ with respect to the class C,if for all other $e^* \in C \backslash \{e\}$ there exists a set $X \in K$ with $e^*(X) \nsubseteq e(X)$. in words:

Definition

let C be a class of estimators (e.g. all estimators $e : 2^D \longrightarrow 2^P$ with a coveringprobability not lower than $1 - \alpha$).
An estimator $e \in C$ is called undominated on $K \subseteq 2^D$ with respect to the class C,if for all other $e^* \in C \backslash \{e\}$ there exists a set $X \in K$ with $e^*(X) \nsubseteq e(X)$. in words: e is undominated on K, if there exists no other estimator in C, that is equally better than e (on K).

Definition

let C be a class of estimators (e.g. all estimators $e : 2^D \longrightarrow 2^P$ with a coveringprobability not lower than $1 - \alpha$).
An estimator $e \in C$ is called undominated on $K \subseteq 2^D$ with respect to the class C,if for all other $e^* \in C \backslash \{e\}$ there exists a set $X \in K$ with $e^*(X) \nsubseteq e(X)$. in words: e is undominated on K, if there exists no other estimator in C, that is equally better than e (on K).

Lemma

Definition

let C be a class of estimators (e.g. all estimators $e : 2^D \longrightarrow 2^P$ with a coveringprobability not lower than $1 - \alpha$).
An estimator $e \in C$ is called undominated on $K \subseteq 2^D$ with respect to the class C,if for all other $e^* \in C \backslash \{e\}$ there exists a set $X \in K$ with $e^*(X) \not\subseteq e(X)$. in words: e is undominated on K, if there exists no other estimator in C, that is equally better than e (on K).

Lemma

Let $e : 2^D \longrightarrow 2^P$ be a compositional estimator.

Definition

let C be a class of estimators (e.g. all estimators $e : 2^D \longrightarrow 2^P$ with a coveringprobability not lower than $1 - \alpha$).
An estimator $e \in C$ is called undominated on $K \subseteq 2^D$ with respect to the class C,if for all other $e^* \in C \backslash \{e\}$ there exists a set $X \in K$ with $e^*(X) \nsubseteq e(X)$. in words: e is undominated on K, if there exists no other estimator in C, that is equally better than e (on K).

Lemma

Let $e : 2^D \longrightarrow 2^P$ be a compositional estimator.If e is undominated on the set of all one-point-sets $\binom{D}{1} = \{\{x\}|x \in D\}$

Definition

let C be a class of estimators (e.g. all estimators $e : 2^D \longrightarrow 2^P$ with a coveringprobability not lower than $1 - \alpha$).
An estimator $e \in C$ is called undominated on $K \subseteq 2^D$ with respect to the class C,if for all other $e^* \in C \backslash \{e\}$ there exists a set $X \in K$ with $e^*(X) \not\subseteq e(X)$. in words: e is undominated on K, if there exists no other estimator in C, that is equally better than e (on K).

Lemma

Let $e : 2^D \longrightarrow 2^P$ be a compositional estimator.If e is undominated on the set of all one-point-sets $\binom{D}{1} = \{\{x\} | x \in D\}$with respect to a class C of <u>set-monotone</u> estimators,

Definition

let C be a class of estimators (e.g. all estimators $e : 2^D \longrightarrow 2^P$ with a coveringprobability not lower than $1-\alpha$).
An estimator $e \in C$ is called undominated on $K \subseteq 2^D$ with respect to the class C,if for all other $e^* \in C \backslash \{e\}$ there exists a set $X \in K$ with $e^*(X) \not\subseteq e(X)$. in words: e is undominated on K, if there exists no other estimator in C, that is equally better than e (on K).

Lemma

Let $e : 2^D \longrightarrow 2^P$ be a compositional estimator.If e is undominated on the set of all one-point-sets $\binom{D}{1} = \{\{x\}|x \in D\}$with respect to a class C of set-monotone estimators, then e is also undominated on 2^D.

Proof.

Let e^* be equally better than e (on 2^D).

Proof.

Let e^* be equally better than e (on 2^D).

$$\exists X \in 2^D \quad : \quad e^*(X) \subsetneq e(X)$$

Proof.

Let e^* be equally better than e (on 2^D).

$$\begin{aligned} \exists X \in 2^D &: \quad e^*(X) \subsetneq e(X) \\ \exists p &\in \quad e(X) \backslash e^*(X) \end{aligned}$$

Proof.

Let e^* be equally better than e (on 2^D).

$$\begin{aligned} \exists X \in 2^D &: & e^*(X) \subsetneq e(X) \\ \exists p &\in & e(X) \backslash e^*(X) \\ \exists x \in X &: & p \in e(\{x\}) \end{aligned}$$

Proof.

Let e^* be equally better than e (on 2^D).

$$\begin{aligned} \exists X \in 2^D &\quad : \quad e^*(X) \subsetneq e(X) \\ \exists p &\quad \in \quad e(X) \backslash e^*(X) \\ \exists x \in X &\quad : \quad p \in e(\{x\}) \end{aligned}$$

from $e^*(\{x\}) \subseteq e^*(X)$ it follows

Proof.

Let e^* be equally better than e (on 2^D).

$$\begin{aligned} \exists X \in 2^D &: e^*(X) \subsetneq e(X) \\ \exists p &\in e(X) \setminus e^*(X) \\ \exists x \in X &: p \in e(\{x\}) \\ \text{from } e^*(\{x\}) \subseteq e^*(X) \text{ it follows} \\ p &\notin e^*(\{x\}) \end{aligned}$$

Proof.

Let e^* be equally better than e (on 2^D).

$$\begin{aligned}
\exists X \in 2^D \quad &: \quad e^*(X) \subsetneq e(X) \\
\exists p \quad &\in \quad e(X) \backslash e^*(X) \\
\exists x \in X \quad &: \quad p \in e(\{x\}) \\
\text{from } e^*(\{x\}) \subseteq e^*(X) \text{ it follows} & \\
p \quad &\notin \quad e^*(\{x\}) \\
e * (\{x\}) \quad &\subsetneq \quad e(\{x\})
\end{aligned}$$

Proof.

Let e^* be equally better than e (on 2^D).

$$
\begin{aligned}
\exists X \in 2^D \quad &: \quad e^*(X) \subsetneq e(X) \\
\exists p \quad &\in \quad e(X) \setminus e^*(X) \\
\exists x \in X \quad &: \quad p \in e(\{x\}) \\
\text{from } e^*(\{x\}) \subseteq e^*(X) \text{ it follows} & \\
p \quad &\notin \quad e^*(\{x\}) \\
e * (\{x\}) \quad &\subsetneq \quad e(\{x\})
\end{aligned}
$$

$\Longrightarrow$ e could not be undominated on $\binom{D}{1}$. □

Another nice implication of set-monotonicity:

Another nice implication of set-monotonicity:

Applicability to fuzzy-set-estimation:

Another nice implication of set-monotonicity:

Applicability to fuzzy-set-estimation:

A fuzzy set defined by its membership-function

Another nice implication of set-monotonicity:

Applicability to fuzzy-set-estimation:

A fuzzy set defined by its membership-function

$\mu_A : D \longrightarrow [0, 1]$

Another nice implication of set-monotonicity:

Applicability to fuzzy-set-estimation:

A fuzzy set defined by its membership-function

$\mu_A : D \longrightarrow [0,1]$

or equivalently by its α*-cut-representation:*

Another nice implication of set-monotonicity:

Applicability to fuzzy-set-estimation:

A fuzzy set defined by its membership-function

$\mu_A : D \longrightarrow [0, 1]$

or equivalently by its α-cut-representation:

α-cut:

Another nice implication of set-monotonicity:

Applicability to fuzzy-set-estimation:

A fuzzy set defined by its membership-function

$\mu_A : D \longrightarrow [0, 1]$

or equivalently by its α-cut-representation:

α-cut:

$A_\alpha := \{d \in D | \mu_A(d) \geq \alpha\}.$

Another nice implication of set-monotonicity:

Applicability to fuzzy-set-estimation:

A fuzzy set defined by its membership-function

$\mu_A : D \longrightarrow [0, 1]$

or equivalently by its α-cut-representation:

α-cut:

$A_\alpha := \{d \in D | \mu_A(d) \geq \alpha\}$.

α-cut-representation:

Another nice implication of set-monotonicity:

Applicability to fuzzy-set-estimation:

A fuzzy set defined by its membership-function

$\mu_A : D \longrightarrow [0,1]$

or equivalently by its α-cut-representation:

α-cut:

$A_\alpha := \{d \in D | \mu_A(d) \geq \alpha\}$.

α-cut-representation:

$A' : [0,1] \longrightarrow 2^D : \alpha \mapsto A_\alpha$

Another nice implication of set-monotonicity:

Applicability to fuzzy-set-estimation:

A fuzzy set defined by its membership-function

$\mu_A : D \longrightarrow [0, 1]$

or equivalently by its α-cut-representation:

α-cut:

$A_\alpha := \{d \in D | \mu_A(d) \geq \alpha\}$.

α-cut-representation:

$A' : [0, 1] \longrightarrow 2^D : \alpha \mapsto A_\alpha$

this allows definition of fuzzy-estimator in terms of the α-cut representation:

Another nice implication of set-monotonicity:

Applicability to fuzzy-set-estimation:
A fuzzy set defined by its membership-function
$\mu_A : D \longrightarrow [0,1]$
or equivalently by its α-cut-representation:
α-cut:
$A_\alpha := \{d \in D | \mu_A(d) \geq \alpha\}$.
α-cut-representation:
$A' : [0,1] \longrightarrow 2^D : \alpha \mapsto A_\alpha$
this allows definition of fuzzy-estimator in terms of the α-cut representation:
$e'(A) : [0,1] \longrightarrow 2^P : \alpha \mapsto e(A_\alpha)$.

Observation: all α-cuts of a fuzzy set are building a chain with respect to the set-inclusion

Observation: all α-cuts of a fuzzy set are building a chain with respect to the set-inclusion $A_{\alpha_1} \subseteq A_{\alpha_2}$ or $A_{\alpha_1} \supseteq A_{\alpha_2}$.

Observation: all α-cuts of a fuzzy set are building a chain with respect to the set-inclusion $A_{\alpha_1} \subseteq A_{\alpha_2}$ or $A_{\alpha_1} \supseteq A_{\alpha_2}$.
So if the estimator is not set-monotone, the constructed fuzzy set $e'(X)$ is in general no well defined fuzzy set.