Measurement Theory: how to tell the truth with numbers

Alexis Tsoukiàs

LAMSADE - CNRS, PSL, Université Paris-Dauphine
tsoukias@lamsade.dauphine.fr

Paris 17,24,22,29/09,01/10/2020

Outline

What the numbers say?

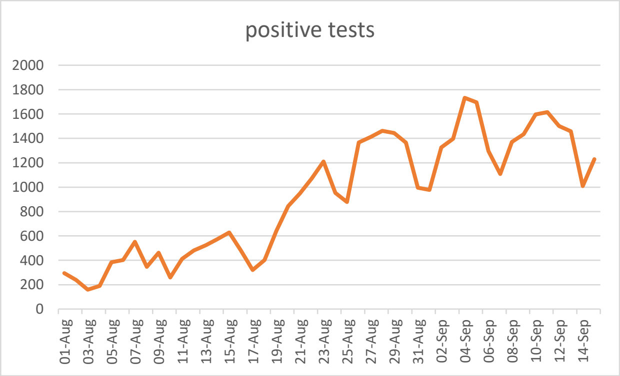


Figure 1: Daily positive tests over time

What the numbers say?

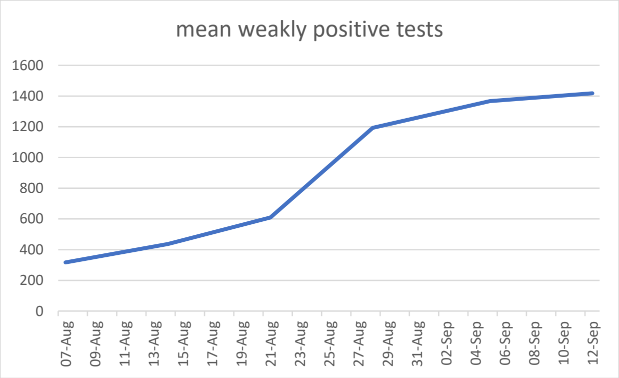


Figure 2: Weakly mean of positive tests over time

What the numbers say?

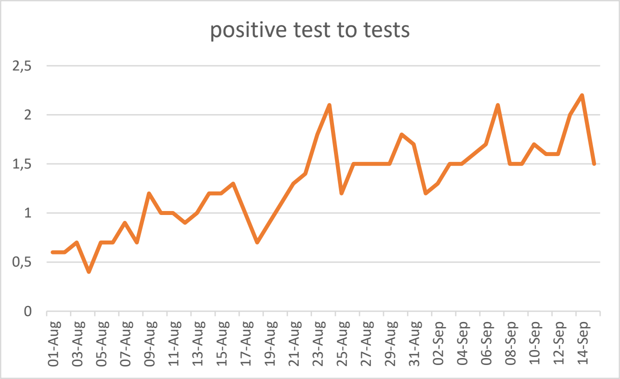


Figure 3: Ratio of positive tests to all tests over time

What the numbers say?

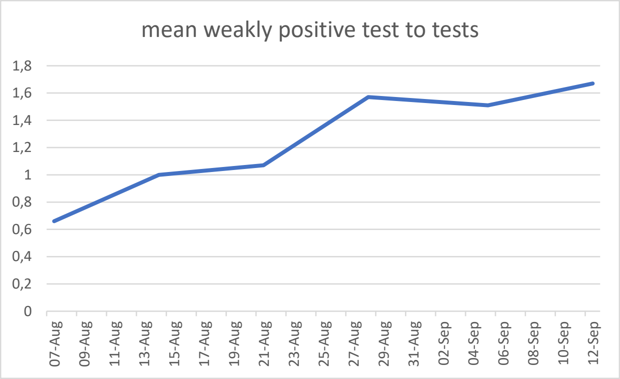


Figure 4: Weakly mean Ratio of positive tests to all tests over time

What the numbers say?

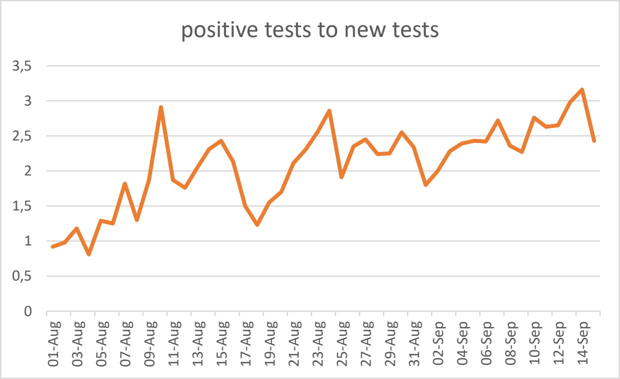


Figure 5: Ratio of positive tests to new tests over time

What the numbers say?

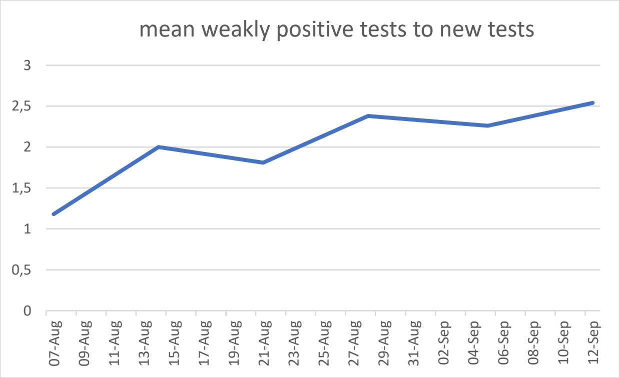


Figure 6: Weakly mean Ratio of positive tests to all tests over time

The Air Quality index: ATMO

We consider 4 pollutants: CO_2, SO_2, O_3 and dust, we observe their concentration (in 1 m^3) and these observations are translated to a scale 1 to 10; 1 being very good and 10 being very bad (for human health). The the ATMO is the worst among the four.

The Air Quality index: ATMO

We consider 4 pollutants: CO_2, SO_2, O_3 and dust, we observe their concentration (in 1 m^3) and these observations are translated to a scale 1 to 10; 1 being very good and 10 being very bad (for human health). The the ATMO is the worst among the four.

Questions:

- Is 5, 5 times worst than 1?
- Is 10, 2 times worst than 5?
- What quantities are precisely compared?
- What is the purpose of this index?
- Can we use it for other purposes?

The Air Quality index

pollutant	CO_2	SO_2	O_3	dust
t_1	3	3	8	8
t_2	1	3	8	2
t_3	7	7	7	7

The Air Quality index

pollutant	CO_2	SO_2	O_3	dust
t_1	3	3	8	8
t_2	1	3	8	2
t_3	7	7	7	7

For the ATMO index t_3 is better than t_2. Is this legitimating (and legitimated)?

Prices of potatoes

The price of potatoes, today at the market near my house is k euros per kg

- What this price represents precisely?
- What does this price reveals?
- Does this price has a general validity?
- Do all consumers consider this price in the same way?
- Can we use this price for other purposes besides actually paying the potatoes?

Elementary Concepts

- Empirical observation: unmediated experience of occurrences in the real world.
- Data: sets, collections of sets and temporal series of empirical observations.
- Information: data used for and within a decision process (purposeful data).
- Knowledge: Information, Past experiences, Culture, Values, Behaviours, Attitudes of a decision maker.

Elementary Concepts

- Decision Process: the set of activities conducted by an individual or an organisation (a decision maker) in designing, choosing and implementing an action.
- Decision Aiding Process: the set of activities occurring between two individual and/or organisations (the client and the analyst) aiming at helping (the client) to follow a decision process.

Our perspective

- Measurement is a necessary activity in order to conduct a decision aiding process. From our perspective this should satisfy 3 requirements.
  - being formal;
  - introduce a dimension of rationality;
  - respect simple scientific standards and norms.

Our perspective

- Measurement is a necessary activity in order to conduct a decision aiding process. From our perspective this should satisfy 3 requirements.
  - being formal;
  - introduce a dimension of rationality;
  - respect simple scientific standards and norms.

Binary relations

Consider A a set

$A = \{x, y, z, w, \cdots\}$

Consider B being a binary relation upon the set A

$B \subseteq A \times A$

$B = \{(x, z), (y, w), (z, w) \cdots\}$

Representations

	x	y	z	w
x	0	0	0	1
y	0	0	1	1
z	0	1	0	1
w	0	1	1	0

Table 1: A binary relation

Representations

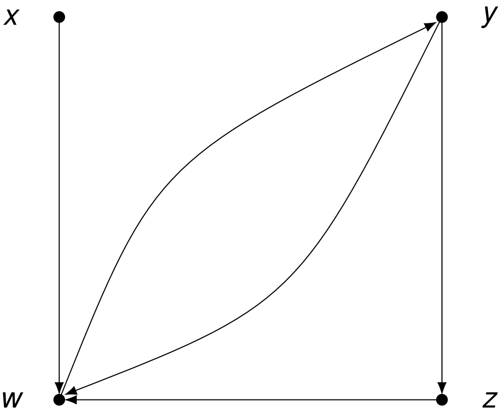


Figure 7: A binary relation

Properties of binary relations

- Reflexive: $\forall x \in A : R(x, x)$
- Irreflexive: $\forall x \in A : \neg R(x, x)$
- Symmetric: $\forall x, y \in A : R(x, y) \rightarrow R(y, x)$
- Asymmetric: $\forall x, y \in A : R(x, y) \rightarrow \neg R(y, x)$
- Complete: $\forall x, y \in A : R(x, y) \vee R(y, x)$
- Transitive: $\forall x, y, z \in A : R(x, y) \wedge R(y, z) \rightarrow R(x, z)$
- Ferrers: $\forall x, y, z, w \in A : R(x, y) \wedge R(z, w) \rightarrow R(x, w) \vee R(z, y)$

Order structures

- Total order: asymmetric, complete and transitive
- Equivalence: symmetric, reflexive and transitive
- Partial order: reflexive and transitive
- Weak order: reflexive, complete and transitive
- Interval order: complete and Ferrers

Semantics

$x \succeq y$

x is at least as good as y

Any order can be seen as the union of two or more binary relations

$\succeq = \succ \cup \sim$

$\succ$ being irreflexive and asymmetric;

$\sim$ being reflexive and symmetric

Generalisation

Generally

$\succeq = \succ_1 \cup \cdots \succ_n \cup \sim_1 \cup \cdots \sim_m$

Relative and Absolute comparisons

Order structures are useful for relative comparisons on any possible set

We can also define “absolute comparisons” between a set A and a set N of norms
$\succeq \subseteq A \times N \cup N \times A$

Numerical Representations

Given a set A and a binary relation $\succeq \subseteq A \times A$

$\exists f : A \mapsto \mathbb{R} \;\; x \succeq y \Leftrightarrow f(x) \geq f(y)$

iff

$\succeq$ is a weak order

Numerical Representations

Given a set A and a binary relation $\succeq \subseteq A \times A$

$\exists l, r : A \mapsto \mathbb{R}, l(x) < r(x) \;\; x \succeq y \Leftrightarrow r(x) \geq l(y)$
iff
$\succeq$ is an interval order
$(x \succ y \Leftrightarrow l(x) > r(y))$

Practically

Practically we use the order among the reals as a substitute of the order among the elements of A

This is only ordinal information. The numbers carry no "quantitative" information. Any monotone increasing transformation would be acceptable.

More complex comparisons

Get $A = \{x, y, z, w \cdots\}$ and then

consider $2^A = \{\{x\}, \{x, y\}, \{x, y, z\}, \cdots\}$ the set of all subsets of A.

We can establish the binary relation $\succcurlyeq \subseteq 2^A \times 2^A$

comparing subsets of A instead of elements of A.

More complex comparisons

Get $A = \{x, y, z, w \cdots\}$ and then

consider $\succeq A \times A$ “at least as good as”

We can establish the binary relation $\succsim \subseteq (A \times A) \times (A \times A)$ such that:

$s(x, y) \succsim s(z, w)$: the ratio(distance) between x and y is at least as good as the one between z and w.

Multiple dimensions

If we have multiple dimensions on which we compare elements of A we get

$\succeq_1 \subseteq A \times A$ and $\succeq_2 \subseteq A \times A$

We can now establish the binary relations $\succsim_1$ and $\succsim_2$ $\subseteq (A \times A) \times (A \times A)$

and from this establish $\succsim_{12} \subseteq (A \times A) \times (A \times A)$ such that $s_1(x, y) \succsim s_2(z, w)$

How to measure a length?

3 Steps

- Compare objects.
- Create and compare new objects.
- Create standard sequences.

Comparing objects: rods

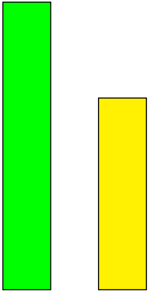


Figure 8: Comparing x to y

Comparing objects: rods

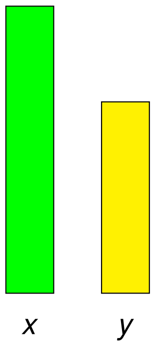


Figure 8: Comparing x to y

Comparing objects: rods

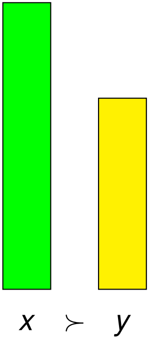


Figure 8: Comparing x to y

Comparing objects: rods

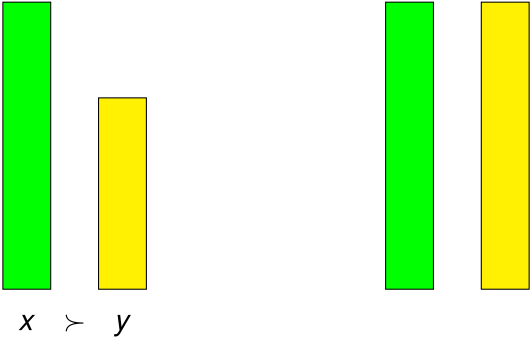


Figure 8: Comparing x to y

Comparing objects: rods

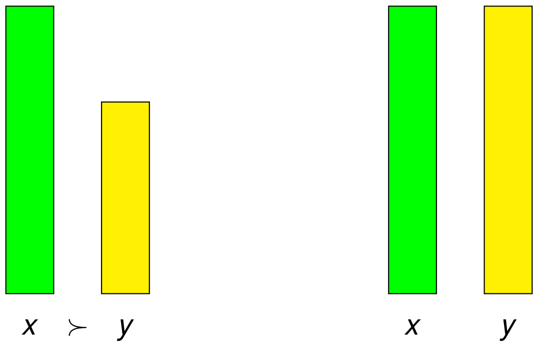


Figure 8: Comparing x to y

Comparing objects: rods

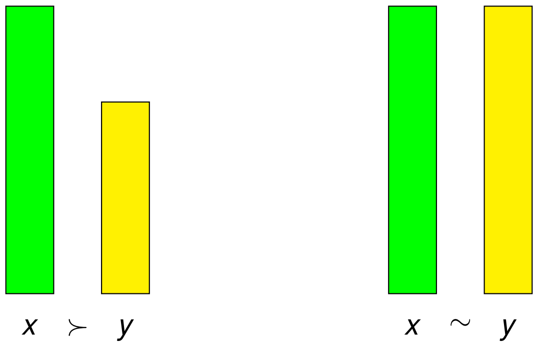


Figure 8: Comparing x to y

What do we get?

$x \succ y$: the extreme of x is higher that the one of y
$x \sim y$: the extreme of x is as high as the one of y

- $x \succ y$ or $y \succ x$ or $x \sim y$.
- $\succ$ is asymmetric and transitive.
- $\sim$ is symmetric and transitive.
- $x \succ y \ \wedge \ y \sim z \ \rightarrow \ x \succ z$
- $x \sim y \ \wedge \ y \succ z \ \rightarrow \ x \succ z$

Results

The relation $\succeq = \succ \cup \sim$ is a weak order (complete and transitive)

Thus:

$$\exists f : A \mapsto \mathbb{R} \text{ such that}$$

$$x \succ y \leftrightarrow f(x) > f(y) \text{ and } x \sim y \leftrightarrow f(x) = f(y)$$

But this is an ordinal measure.

Creating and Comparing objects: rods

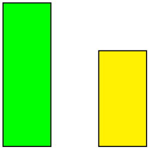


Figure 9: Comparing x to y

Creating and Comparing objects: rods

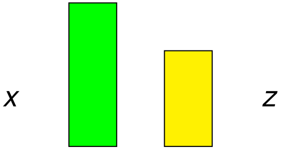


Figure 9: Comparing x to y

Creating and Comparing objects: rods

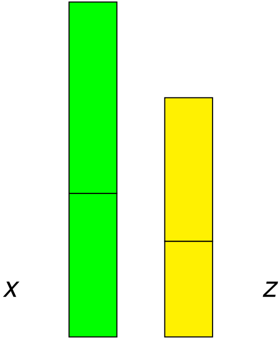


Figure 9: Comparing x to y

Creating and Comparing objects: rods

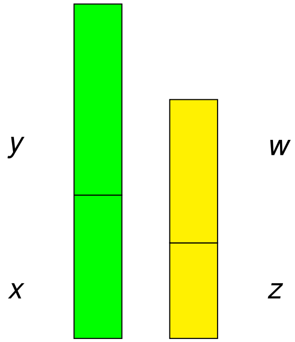


Figure 9: Comparing x to y

Creating and Comparing objects: rods

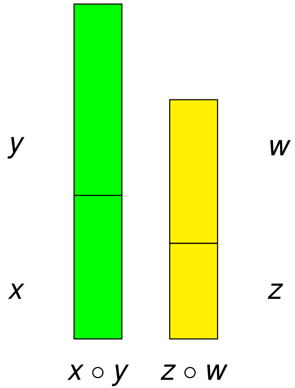
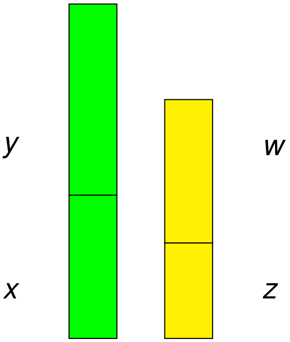


Figure 9: Comparing x to y

Creating and Comparing objects: rods



$x \circ y \succ z \circ w$

Figure 9: Comparing x to y

What are we looking for?

We need a function f such that:

$$f(x \circ y) = f(x) + f(y)$$

$$x \succ y \wedge z \sim w \rightarrow x \circ z \succ y \circ w$$

More comparisons

Consider 5 objects x, y, z, w, t

and the following empirical observations:

$$x \circ t \succ z \circ w \succ x \circ y \succ t \succ w \succ z \succ y \succ x$$

More numerical representations

Consider now the following numerical representations:

	L_1	L_2	L_3
x	14	10	14
y	15	91	16
z	20	92	17
w	21	93	18
t	28	99	29

L_1, L_2 and L_3 capture the simple order among α_{1-5}, but L_2 fails to capture the order among the combinations of objects.

BUT ...

If we compare new combinations

For L_1: $y \circ z \sim x \circ w$

For L_3: $y \circ z \succ x \circ w$

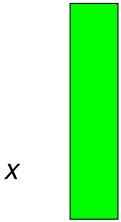
BUT ...

If we compare new combinations

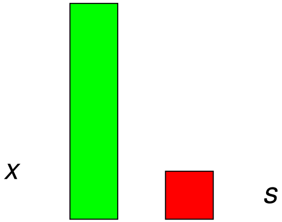
For L_1: $y \circ z \sim x \circ w$
For L_3: $y \circ z \succ x \circ w$

We need to establish how sequences are measured.

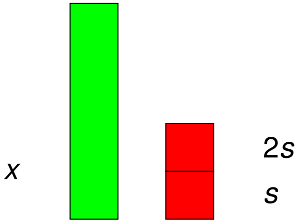
Creating standard sequences



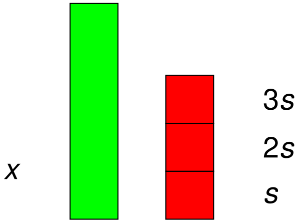
Creating standard sequences



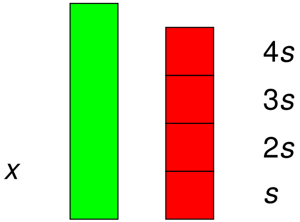
Creating standard sequences



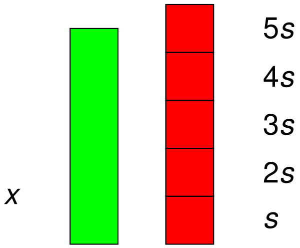
Creating standard sequences



Creating standard sequences



Creating standard sequences



Creating standard sequences

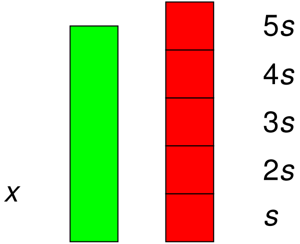


Figure 10: $5s \succ x \succ 4s$: If $f(s) = 1$ then $5 > f(x) > 4$

How to converge?

First option

Create smaller standards
repeat until for some s you obtain $x \sim s \circ s \circ s \cdots s$ k times.
Then $f(x) = kf(s)$.

How to converge?

First option

Create smaller standards
repeat until for some s you obtain $x \sim s \circ s \circ s \cdots s$ k times.
Then $f(x) = kf(s)$.

Second option

Make copies of the object
Concatenate the copies
Repeat until $x \circ x \circ \cdots x$ n times $\sim s \circ s \circ \cdots s$ m times
Then $nf(x) = mf(s)$.

What do we get?

Any object is compared to a sequence of standards/units

Independently from what the standard/unit is, the ratio of units between any two objects remains constant.

What do we get?

Any object is compared to a sequence of standards/units

Independently from what the standard/unit is, the ratio of units between any two objects remains constant.

This is a ratio scale.

New setting

Consider the following case:

you observe a system at time t_1 and then the same system at time t_2 and you want to represent the difference between the two states.

New setting

Consider the following case:

you observe a system at time t_1 and then the same system at time t_2 and you want to represent the difference between the two states.

What is standard here?

Temperature

You need to establish an origin with respect to which measure differences

You then arbitrary establish as standard a state of the system you can objectively observe: icing! The temperature where the water ices: t_i.

Temperature

You need to establish an origin with respect to which measure differences

You then arbitrary establish as standard a state of the system you can objectively observe: icing! The temperature where the water ices: t_i.

This is the origin of your measures

Temperature

But we also need a standard sequence for the distance

We establish another objectively observable state: boiling. The temperature where the water boils: t_b.

Temperature

But we also need a standard sequence for the distance

We establish another objectively observable state: boiling. The temperature where the water boils: t_b.

The standard $\frac{t_b - t_i}{100}$ is your unit

What do we get?

Any state is compared to a sequence of standards/units, with respect a common origin

Independently from the origin and the unit, the ratio among the differences of units remains constant.

What do we get?

Any state is compared to a sequence of standards/units, with respect a common origin

Independently from the origin and the unit, the ratio among the differences of units remains constant.

This is an interval scale.

Summarising

Ordinal scales

Any monotone non decreasing transformation of the scale delivers the same information with respect to the empirical observations and the order between them.

Summarising

Ordinal scales

Any monotone non decreasing transformation of the scale delivers the same information with respect to the empirical observations and the order between them.

Ratio scales

Any proportional transformation (of the type αx) of the scale delivers the same information with respect to the empirical observations and the order between them.

Summarising

Ordinal scales

Any monotone non decreasing transformation of the scale delivers the same information with respect to the empirical observations and the order between them.

Ratio scales

Any proportional transformation (of the type αx) of the scale delivers the same information with respect to the empirical observations and the order between them.

Interval scales

Any affine transformation (of the type $\alpha x + \beta$) of the scale delivers the same information with respect to the empirical observations and the order between them.

What about psychometrics?

Factor analysis (or other multivariate statistics) upon multidimensional responses to psychological stimuli.

- Irregular scales
- Observation bias
- Self confirming bias

Meaningfulness 1

Each type of scale allows a certain type of admissible transformations which do not modify arbitrarily the empirical observations

Ordinal $\quad f(x) \geq f(y) \quad \Leftrightarrow \quad \Phi(f(x)) \geq \Phi(f(y))$

.

Meaningfulness 1

Each type of scale allows a certain type of admissible transformations which do not modify arbitrarily the empirical observations

Ordinal	$f(x) \geq f(y)$	$\Leftrightarrow$	$\Phi(f(x)) \geq \Phi(f(y))$
Ratio	$\frac{f(x)}{f(y)} = k$	$\Leftrightarrow$	$\frac{\Phi(f(x))}{\Phi(f(y))} = k$

.

Meaningfulness 1

Each type of scale allows a certain type of admissible transformations which do not modify arbitrarily the empirical observations

.

Ordinal	$f(x) \geq f(y)$	$\Leftrightarrow$	$\Phi(f(x)) \geq \Phi(f(y))$
Ratio	$\frac{f(x)}{f(y)} = k$	$\Leftrightarrow$	$\frac{\Phi(f(x))}{\Phi(f(y))} = k$
Interval	$\frac{f(x)-f(y)}{f(z)-f(w)} = k$	$\Leftrightarrow$	$\frac{\Phi(f(x))-\Phi(f(y))}{\Phi(f(z))-\Phi(f(w))} = k$

Meaningfulness 2

Manipulations

Any type of manipulation of numerical information needs to respect the admissible transformations property. In this case we talk about meaningful manipulations.

Example

$x \succ y \succ z \succ w \succ t$

$$f(t) \quad f(w) \quad f(z) \quad f(y) \quad f(x)$$

Example

$x \succ y \succ z \succ w \succ t$

	$f(t)$	$f(w)$	$f(z)$	$f(y)$	$f(x)$
L1	1	2	3	4	5

Example

$x \succ y \succ z \succ w \succ t$

	$f(t)$	$f(w)$	$f(z)$	$f(y)$	$f(x)$
L1	1	2	3	4	5
L2	1	2	3	4	10

Example

$x \succ y \succ z \succ w \succ t$

	$f(t)$	$f(w)$	$f(z)$	$f(y)$	$f(x)$
L1	1	2	3	4	5
L2	1	2	3	4	10

The mean of L1 is 3 and the mean object is z. The mean of L2 is 4 and the mean object is y. The median instead is always z independently from any numerical coding of the ordinal scale.

A mean is meaningless in presence of ordinal information, while a median is meaningful.

More about manipulations

alternatives	g_1	g_2
h	2000	500
a	160	435
b	400	370
c	640	305
d	880	240
e	1120	175
f	1360	110
g	1600	45

Table 2: Weighted sum

More about manipulations

alternatives	g_1	g_2	g_1^n	g_2^n
h	2000	500	100.00	100.00
a	160	435	8.00	87.00
b	400	370	20.00	74.00
c	640	305	32.00	61.00
d	880	240	44.00	48.00
e	1120	175	56.00	35.00
f	1360	110	68.00	22.00
g	1600	45	80.00	9.00

Table 2: Weighted sum

More about manipulations

alternatives	g_1	g_2	g_1^n	g_2^n	Score	Rank
h	2000	500	100.00	100.00	100.0	1
a	160	435	8.00	87.00	47.5	2
b	400	370	20.00	74.00	47.0	3
c	640	305	32.00	61.00	46.5	4
d	880	240	44.00	48.00	46.0	5
e	1120	175	56.00	35.00	45.5	6
f	1360	110	68.00	22.00	45.0	7
g	1600	45	80.00	9.00	44.5	8

Table 2: Weighted sum

More about manipulations

alternatives	g_1	g_2	g_1^n	g_2^n	Score	Rank
h	2000	**700**	100.00	100.00		
a	160	435	8.00	62.14		
b	400	370	20.00	52.86		
c	640	305	32.00	43.57		
d	880	240	44.00	34.29		
e	1120	175	56.00	25.00		
f	1360	110	68.00	15.71		
g	1600	45	80.00	6.43		

Table 3: Weighted sum

More about manipulations

alternatives	g_1	g_2	g_1^n	g_2^n	Score	Rank
h	2000	**700**	100.00	100.00	100.0	1
a	160	435	8.00	62.14	35.07	**8**
b	400	370	20.00	52.86	36.43	**7**
c	640	305	32.00	43.57	37.79	**6**
d	880	240	44.00	34.29	39.14	**5**
e	1120	175	56.00	25.00	40.50	**4**
f	1360	110	68.00	15.71	41.86	**3**
g	1600	45	80.00	6.43	43.21	**2**

Table 3: Weighted sum

More about manipulations

alternatives	g_1	g_2	g_1^n	g_2^n	Score	Rank
h	2000	**700**	100.00	100.00	100.0	1
a	**165**	**450**	8.25	64.29	36.27	**8**
b	400	370	20.00	52.86	36.43	**7**
c	640	305	32.00	43.57	37.79	**6**
d	880	240	44.00	34.29	39.14	**5**
e	1120	175	56.00	25.00	40.50	**4**
f	1360	110	68.00	15.71	41.86	**3**
g	1600	45	80.00	6.43	43.21	**2**

Table 4: Weighted sum

More about manipulations

alternatives	g_1	g_2	g_1^n	g_2^n	Score	Rank
h	2000	**700**	100.00	100.00	100.0	1
a	**165**	**450**	8.25	64.29	36.27	**8**
b	400	370	20.00	52.86	36.43	**7**
c	640	305	32.00	43.57	37.79	**6**
d	880	240	44.00	34.29	39.14	**5**
e	1120	175	56.00	25.00	40.50	**4**
f	1360	110	68.00	15.71	41.86	**3**
g	1600	45	80.00	6.43	43.21	**2**

Table 4: Weighted sum

J.-Ch. Billaut, D. Bouyssou, Ph. Vincke, “Should you believe the Shanghai index?” *Scientometrics*, vol. 84, 237 - 263, 2010.

Usefulness 1

Consider $f(x_1)\cdots f(x_n)$ the ratio scale measurement of the set $A = \{x_1 \cdots x_n\}$. Then if $\Phi(f(x_j)) = \alpha f(x_j)$ is an admissible transformation

$$\text{IF } \frac{\sum_{j=1}^{n} f(x_j)}{n} = k \text{ THEN } \frac{\sum_{j=1}^{n} \Phi(f(x_j))}{n} = \alpha k$$

$$\text{IF } (\prod_{j=1}^{n} f(x_j))^{1/n} = k \text{ THEN } (\prod_{j=1}^{n} \Phi(f(x_j)))^{1/n} = \alpha k$$

Usefulness 2

Both the arithmetic mean and the geometric mean are meaningful manipulations of ratio scales measurement.

But if $f(x_j)$ are the measures of the edges of a solid, the arithmetic mean tells us nothing, while the geometric mean tells us something about the volume of the solid.

Usefulness 2

Both the arithmetic mean and the geometric mean are meaningful manipulations of ratio scales measurement.

But if $f(x_j)$ are the measures of the edges of a solid, the arithmetic mean tells us nothing, while the geometric mean tells us something about the volume of the solid.

Not all meaningful manipulations are useful

Legitimacy

- Racial and Religious Statistics
- Cultural misunderstandings and misinterpretations
- Historical biases

Legitimacy

- Racial and Religious Statistics
- Cultural misunderstandings and misinterpretations
- Historical biases

Not all meaningful and useful manipulations are legitimated

What is a statistic?

Describe a population through a number of features in a synthetic way, summarising it

- We need to establish who the population is.
- We need to establish which are the relevant features.
- We need empirical observations.

Descriptive Statistics

Describe a population using summarising indexes such as means, medians, deviations etc.

- What is the average age of the voters for X?
- What is the median age of positively tested Covid-19 patients recovered in ICU?
- How is the population clustered following residence and age?

Inferential Statistics

Try to infer the likelihood of events out of empirical observations upon a given population

- What is the likelihood of X being elected given the age distribution?
- What is the likelihood of getting seriously ill once positively tested?
- How do we expect the behaviour of the population to evolve?

Two ways to make a statistic

Observe the whole population and describe one or more features

Observe a sample of the population and describe one or more features

Two ways to make a statistic

Observe the whole population and describe one or more features

Observe a sample of the population and describe one or more features

NB. Samples need to be representative of the population

Empirical observations which do not observe the whole population or a representative sample ARE NOT statistics

Examples

- Testing 1 out of any 100 lamps provides a statistic? Of what?
- Are the daily results of the COVID-19 tests a statistic? Of what?
- Which empirical observation constitutes a statistic of the COVID-19 disease within the French population?

Examples

- Testing 1 out of any 100 lamps provides a statistic? Of what?
- Are the daily results of the COVID-19 tests a statistic? Of what?
- Which empirical observation constitutes a statistic of the COVID-19 disease within the French population?

Examples

- Testing 1 out of any 100 lamps provides a statistic? Of what?
- Are the daily results of the COVID-19 tests a statistic? Of what?
- Which empirical observation constitutes a statistic of the COVID-19 disease within the French population?

A French statistic

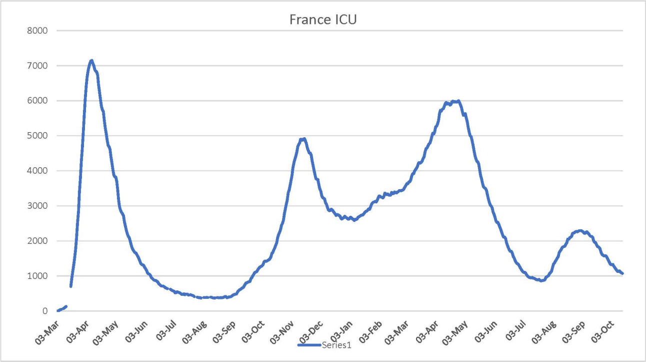


Figure 11: French ICU patients daily

How to perform a statistic?

Direct observation

Observe a population or a sample through one or more variables (features), test one or more hypotheses and observe any correlation among them.

How to perform a statistic?

Direct observation

Observe a population or a sample through one or more variables (features), test one or more hypotheses and observe any correlation among them.

Experimental observation

Observe a population or a sample as before, then modify the parameters of the observation and test again. This could reveal correlations not directly observable.

How to perform a statistic?

Direct observation

Observe a population or a sample through one or more variables (features), test one or more hypotheses and observe any correlation among them.

Experimental observation

Observe a population or a sample as before, then modify the parameters of the observation and test again. This could reveal correlations not directly observable.

NB. Experimental design biases

Some fun

1. More than 98 percent of convicted felons are bread users.
2. Fully HALF of all children who grow up in bread-consuming households score below average on standardised tests.
3. In the 18th century, when virtually all bread was baked in the home, the average life expectancy was less than 50 years;
4. More than 90 percent of violent crimes are committed within 24 hours of eating bread.
5. Primitive tribal societies that have no bread exhibit a low incidence of cancer, Alzheimer's, Parkinson's disease, and osteoporosis.
6. Bread has been proven to be addictive. Subjects deprived of bread and given only water to eat begged for bread after as little as two days.
7. Most American bread eaters are utterly unable to distinguish between significant scientific fact and meaningless statistical babbling.

Correlations

Empirical observations can be correlated and so statistics:

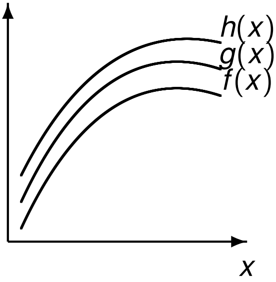


Figure 12: Three correlated frequencies

However ...

We cannot claim

$\forall x\ g(x)$ is a consequence of $f(x)$ and the same among any pair of the frequencies

However ...

We cannot claim

$\forall x \ g(x)$ is a consequence of $f(x)$ and the same among any pair of the frequencies

What are we talking exactly about?

Logical implication: $\alpha \rightarrow \beta$
Material implication: $A \subseteq B$
Causality: A is the cause of B

Different causalities

Logical causality

B is a logical consequence of A

Different causalities

Logical causality

B is a logical consequence of A

Counterfactual causality

If A was not the case then B will never hold.

Different causalities

Logical causality

B is a logical consequence of A

Counterfactual causality

If A was not the case then B will never hold.

Conditional causality

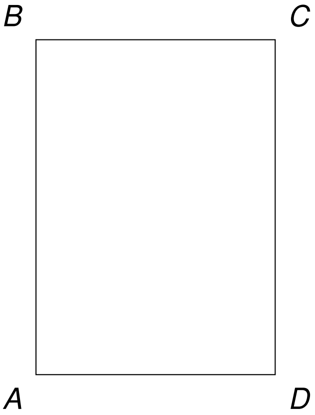
$P(B|A) > P(B)$: the likelihood of B occurring when A is the case is larger than the likelihood of B.

Conditional inferences and likelihoods

Consider disease A with 1% presence among the population

If you take test B which is 99% reliable and it is positive, what is the probability you are indeed an A patient?

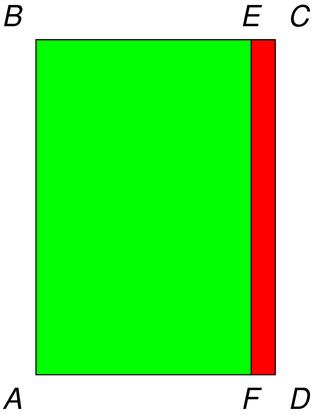
Visual conditionals



ABCD: the population

Figure 13: Conditional Likelihoods

Visual conditionals



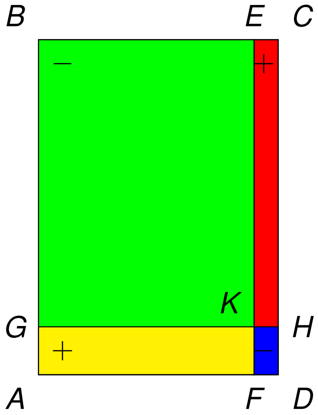
ABCD: the population

ABEF: the healthy

FECD: the sick

Figure 13: Conditional Likelihoods

Visual conditionals



ABCD: the population

ABEF: the healthy

FECD: the sick

AGKF: the false positive

GBEK: the true negative

FKHD: the false negative

KECH: the true positive

Figure 13: Conditional Likelihoods

Conditioning

Suppose you are positively tested

What you really know is that you have a positive test which means you are in one among the two areas: the false positive and the true positive. The probability of being sick is thus, the portion of true positive among all the positive tested

Conditioning

Suppose you are positively tested

What you really know is that you have a positive test which means you are in one among the two areas: the false positive and the true positive. The probability of being sick is thus, the portion of true positive among all the positive tested

$$P(S|+) = \frac{KECH}{KECH + AGKF}$$

Bayes Theorem

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Conclusions

- Do not play with numbers.
- Pay attention to the formal properties number satisfy and need to satisfy.
- Pay attention to the purpose for which numbers are adopted as a model of the reality.
- Pay attention to who uses the numbers, how and where. Do not hesitate to be critical.
- Be honest with you, the clients and the citizens.

References

Two essential readings:
- F. Roberts, *Measurement Theory for the Social Sciences*, Addison Wesley, Boston, 1979, 2nd edition Cambridge University Press, 1985.
- D. Bouyssou, Th. Marchant, M. Pirlot, P. Perny, A. Tsoukiàs, Ph. Vincke, *Evaluation and Decision Models: a critical perspective*, Kluwer Academic, Dordrecht, 2000, (https://www.lamsade.dauphine.fr/~tsoukias/download/books/Book.pdf)

Further reading:
- D.H. Krantz, R.D. Luce, P. Suppes, A. Tversky, *Foundations of measurement*, Academic Press, New York, 1971.
- G. Gigerenzer, *Calculated risks: How to know when numbers deceive you*, Simon and Schuster, New York, 2002.
- J. Michell, *Measurement in Psychology*, Cambridge University Press, Cambridge MA, 1999.
- N. Rescher, *Introduction to Value Theory*, Prentice Hall, Englewood Cliffs, 1969.