



UNI EN ISO 9001:2008
certificato da ICIM
n. 4651/0



Statistica per che cosa? Consulenza in ambito Statistico

Politecnico di Torino, Marzo 2011

Statistica per che cosa? Consulenza in ambito Statistico

It is not easy to identify the border between what is certainly and objectively measurable and what is not. It would seem that everything that can be expressed with numbers or, more generally, that which can be formalized in mathematical models is objectively measurable. On the other hand, what it is not possible to formalize or express through numbers would seem not to be objectively measurable.

Statistica per che cosa? Consulenza in ambito Statistico

Here below I report some difficulties in communication through numbers which I shall summarise in four questions: (1) What do we communicate with numbers? (2) Who builds those numbers? (3) Who communicates the numbers? and (4) To whom do they communicate the numbers?

Statistica per che cosa? Consulenza in ambito Statistico

...I want to communicate instead the feeling coming from years of work in the field of data analysis, the feeling that I have contributed towards building a sand castle, the feeling of disillusion...

...it is worth fighting for a better and more "true" representation of reality because Objectivity is not given, Objectivity is a conquest...

Statistica per che cosa? Consulenza in ambito Statistico

To show as a context may be differently represented with words, it is instructive to read the various descriptions that may be given of a banal event, such as the one which took place on what could be any bus S during the rush hour. This is narrated in 99 different ways by R. Queneau in his book *Exercises in Style*

Statistica per che cosa? Consulenza in ambito Statistico

Notazioni

Sulla S, in un'ora di traffico. Un tipo di circa ventisei anni, cappello floscio con una cordicella al posto del nastro, collo troppo lungo, come se glielo avessero tirato. La gente scende. Il tizio in questione si arrabbia con un vicino. Gli rimprovera di spingerlo ogni volta che passa qualcuno. Tono lamentoso, con pretese di cattiveria. Non appena vede un posto libero, vi si butta. Due ore più tardi lo incontro alla Cour de Rome, davanti alla Gare Saint-Lazare. È con un amico che gli dice: « Dovresti far mettere un bottone in più al soprabito ». Gli fa vedere dove (alla sciancratura) e perchè.

Statistica per che cosa? Consulenza in ambito Statistico

Volgare

Aho! Annavo a magnà e te monto su quer bidone de la Esse - e 'an vedi? - nun me vado a incoccià con 'no stronzo con un collo cche pareva un cacciavite, e 'na trippa sur cappello? E quello un se mette a baccaglià con st'altro burino perchè - dice - je' acciacca er ditone? Te possino! Ma cche voi, ma cchi spinge? e certo che spinge! chi, io? ma va a magnà er sapone!

'Nzomma, meno male che poi se va a sede.

E bastasse! Sarà du' ore dopo, chi s'arrivede? Lo stronzo, ar Colosseo, che sta a complottà con st'altro quà che se crede d'esse er Chistian Dior, er Missoni, che so, er Mister Facis, li mortacci sui! E metti un bottone de quà, e sposta un bottone de là, a acchittate così alla vitina, e ancora un po' ce faceva lo spacchetto, che era tutta 'na froceria che te dico. Ma vaffanculo!

Statistica per che cosa? Consulenza in ambito Statistico

Insiemistica

Nell'autobus S si consideri l'insieme A dei passeggeri seduti e l'insieme D dei passeggeri in piedi. A una fermata data si trovi l'insieme P dei passeggeri in attesa. Sia C l'insieme dei seduti e sia esso un sottoinsieme di P che rappresenti l'unione di C' quale insieme dei passeggeri che restano sulla piattaforma e di C'' quale insieme di coloro che vanno a sedersi. Si dimostri che l'insieme C'' è vuoto.

Sia Z l'insieme dei fricchettoni e $\{z\}$ l'intersezione di Z e C' , ridotto a un solo elemento. A seguito della iniezione dei piedi di z su quelli di y (elemento qualsiasi di C' che sia differente da z) si produce un insieme M di parole emesse da z . L'insieme C'' essendo nel frattempo divenuto non vuoto, dimostrare come esso si componga dell'unico elemento z . Sia ora P' l'insieme dei pedoni che si trovano di fronte alla Gare Saint-Lazare, sia $\{z, z'\}$ l'intersezione di Z e P' , sia B l'insieme dei bottoni di soprabito di z , B' l'insieme delle posizioni possibili di detti bottoni secondo z' : dimostrare che l'iniezione di B in B' non è una bi-iniezione.

Statistica per che cosa? Consulenza in ambito Statistico

Sonetto

Tanto gentile la vettura pare che va da
Controscarpa a Ciampelletto che le genti gioiose
a si pigiare vi van, e va con esse un giovinetto.
Alto ha il collo, e il cappello deve stare avvolto di
un gallone a treccia stretto: potrai tu biasimarlo
se un compare iroso insulta, che gli pigia il retto?
Ora s'è assiso. Sarà d'uopo almeno ritrovarlo al
tramonto, quando poi non lontano da luogo ove
sta il treno s'incontri con l'amico, che gli eroi
della moda gli lodi, e non sia alieno
dall'aumentare li bottoni suoi.

Statistica per che cosa? Consulenza in ambito Statistico

Instead it would seem that reality translated into numbers is in some way more true than that translated into words. That regardless of evidence to the contrary: probably the “true” number of tourists, the “true” number of the employed, the “true” number of the victims of that natural disaster will never be known. As for other cases, we can think of the rate of inflation or the forecasts of the development of the Gross Domestic Product (GDP), or the number of participants at a political or trade union demonstration. The construction process of such data is so complex and at times even biased^[i], that it is difficult to trust in them blindly.

^[ii] That is shaped by particular point of view: “...In short, even official statistics are social products, shaped by the people and organizations that create them...” [Best, J. 2001, p. 26]; “...advocates who conduct their own surveys can decide how to interpret the results..” [Best, J. 2001, p. 48]. Best’s book contains a lot of examples of uses and misuses of numbers to which I refer readers for reference.

Statistica per che cosa? Consulenza in ambito Statistico

The number is usually the result produced by a more or less complex calculation carried out by people who are experts in the techniques required, in particular in statistics. It ranges from the most elementary case, the enumeration of objects, to others that are less elementary, in which the processing of the number in question involves the use of complex computational procedures

Statistica per che cosa? Consulenza in ambito Statistico

In other cases, which are less immediate, the production of the “number” in question involves not only counting or an estimate, but also a forecast or, better, a predictive model. The model is the crystal ball that the experts use to make their prophecies

For example: $Q = a + b_1P + b_2S$

Q=Quantity of refuse per family, P=N. of persons,

S=Residence Surface Area, b1 and b2= estimates

Statistica per che cosa? Consulenza in ambito Statistico

In this section we will see how models give sense to numbers. Models may be viewed as a particular (formalized) point of view, or conceptual framework or context [Paulos, 1998, p. 14*[i]*] in which numbers find their true reference or correspondence to the reality that they describe. Once the reality has been captured in a net of conceptual frameworks or models, numbers begin to be related to it in an unambiguous way and with an unambiguous meaning

[i] "While some figures are almost self-explanatory, statistics without any context always run the risk of being arid, irrelevant, even meaningless."

Statistica per che cosa? Consulenza in ambito Statistico

The study compares school achievement in two different types of high schools, one of which is of a general type, the other one of vocational type [OECD 1999]. The students of the first one show better achievement than the students in the other one. As students with lower SES are more likely to be enrolled in vocational schools there is a strong relationship between SES and achievement, although this may be not true inside a single school type (see also [Raudenbush, S.W., Bryk A.S. 2002, p.16-22]).

Statistica per che cosa? Consulenza in ambito Statistico

Imagine transferring the situation described for two schools into one school and referring the findings related to the observed differences in achievement in two school types to two classrooms. In the first classroom the teacher, Aristogitone, is severe; in the other second classroom the teacher is Valdo, who is less strict (for the sake of simplicity there is only one teacher per classroom).

Statistica per che cosa? Consulenza in ambito Statistico

Both of them are teachers who evaluate the achievement of their students in a strictly technical way and in their classrooms the probability of obtaining good achievements is exactly the same for each level of a student's SES. The only difference is that with Aristogitone the probability to fail is 30%, while with Valdo it is 10%.

Statistica per che cosa? Consulenza in ambito Statistico

Both teachers are also perfectly impartial in their evaluations. But knowing that Aristogitone's students are more likely to come from a lower SES we would argue that they are, *taken together*, biased and unfair with respect to those students...

Both the opinions are well supported, but what is more likely to happen is that teachers will think that their situation represents the true reality and therefore that there is no relationship between SES and achievement [Bonifacio, 1987, p. 98]. This is one of those cases in which true observations give rise to false, unexpected consequences [*effets pervers*, see Boudon 1977].

Statistica per che cosa? Consulenza in ambito Statistico

An analogous example is reported by Paulos [Paulos, J.H, 1998, p. 39]. The example is drawn from race relations in USA and shows how, supposing the same diffusion of racism among white and black people (10 percent of racists in each group), blacks will suffer disproportionately from racism due to the different marginal distribution of whites and blacks in the population[i].

[i] The probability that Paulos and I both gave not only almost the same example, but also used actors with almost the same name, Waldo [Paulos, J.A. 1998, p.91] and Valdo respectively, is very low. In some sense we always have to be careful with numbers, even with small numbers, in spite “of the stunning insignificance of the vast majority of coincidences” [Paulos, 1998, p. 61]

Statistica per che cosa? Consulenza in ambito Statistico

$$Pr(R_1|C_i) = Pr(R_1|E_1) + Pr(E_2|C_i) \cdot [Pr(R_1|E_2) - Pr(R_1|E_1)]$$

where:

$Pr(R_1|C_i)$ = probability to fail R_1 given the SES level C_i

$Pr(R_1|E_1)$ = probability to fail given the school type (General E_1)

$Pr(E_2|C_i)$ = probability to be enrolled in the vocational school given the SES level C_i

$[Pr(R_1|E_2) - Pr(R_1|E_1)]$ = difference between the probabilities to fail between the school types

Statistica per che cosa? Consulenza in ambito Statistico

This view simply tells us that literature and mathematics are both “right” modes to know the world, but in a sort of specialized and complementary way. Reality is not so simple to fit definitely in a model. Or vice versa, a model could not be so complex to fit the reality at a delta level, where delta is less than a given ε taken as small as possible (and also even if it were possible it would be useless)

Statistica per che cosa? Consulenza in ambito Statistico

In order for there to be a *model* and an expert to *assess* it a *problem* arises from an *objective* to be reached

The interaction among the protagonists (*problem, objectives, assessment, model* on the one hand; *customer, experts, families* on the other) constructs the number

Statistica per che cosa? Consulenza in ambito Statistico

The *expert* is a subject of the interaction as he or she knows the process of the construction of the number

The *customer* is the subject of interaction as he or she knows better than the others the objectives and problems

In turn the *subject* of the survey is the subject of interaction as he reacts to stimuli to which he is subjected

Statistica per che cosa? Consulenza in ambito Statistico

I am referring here to another source of misinterpretation of numbers pertaining to the communication world: numbers are not only presented as they are, but transformed in a simpler, more intuitive fashion. This simplification (nearly always graphication) is often an oversimplification in which something important gets lost: the way in which numbers have been built, that is the underlying logic or model

Statistica per che cosa? Consulenza in ambito Statistico



Statistica per che cosa? Consulenza in ambito Statistico



Statistica per che cosa? Consulenza in ambito Statistico



The image shows a screenshot of a news segment from Tg2. The title is "LE MISURE DEGLI UOMINI". The statistics listed are:

- ALTEZZA MEDIA: 1,75 metri
- PESO MEDIO: 77,2 kg
- INDICE DI MASSA CORPOREA: 25,26

The source is cited as "FONTE: METIS RICERCHE". The video player interface at the bottom shows a progress bar at 00:05:16 / 00:16:49 and a thumbs-up icon.

Statistica per che cosa? Consulenza in ambito Statistico



Statistica per che cosa? Consulenza in ambito Statistico



Statistica per che cosa? Consulenza in ambito Statistico



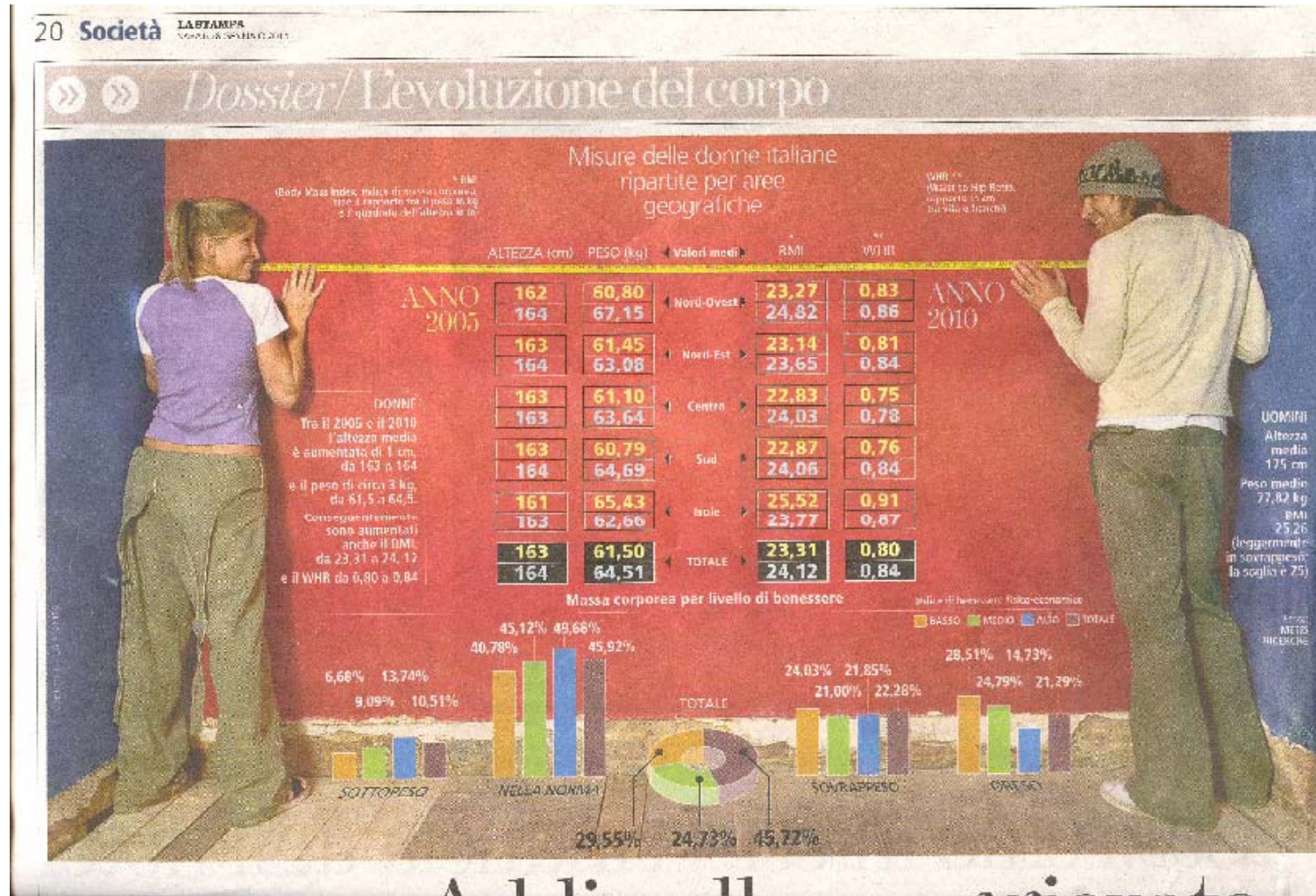
Statistica per che cosa? Consulenza in ambito Statistico

Despite this “intentionality” [PAULOS 1998, p. 92] of models, this plasticity of models, today nearly every opinion or policy maker uses numbers (data) to support his arguments without a model. Or better they put data inside an empty model, that is a model that models nothing.

Statistica per che cosa? Consulenza in ambito Statistico

Despite the fact that they use more data than in the past and therefore they need - even more than in the past - empirical research tools such as opinion poll surveys, they have not developed what I would call an *objective feeling*, where *objective feeling* is the special aptitude to respect numbers and their special ability to represent, in some way, reality. What happens instead is that everyone looks for the most convenient numbers that best apply to a particular, usually his own, point of view

Statistica per che cosa? Consulenza in ambito Statistico



Statistica per che cosa? Consulenza in ambito Statistico

L'italiana ha meno curve

Le indagini antropometriche confermano: fianchi stretti, statura in crescita, e qualche chilo in più

ROSSETTA SALIZZI

Un po' più alte (un centimetro, da 1,63 a 1,64). Un po' più rotonde (ci sono sempre quei tre chili di troppo, si passa da 61,5 a 64,5), ma con meno curve, soprattutto le ragazze. Il modello androgino, da passerella, meno femminile e meno mediterraneo, potrebbe essere una bella rivoluzione, nella patria della magliorata, Sophia Loren in testa, passata per Sabina Ferilli e Luisa Ranieri. Però le top hanno fatto scuola, alto, pectori, fianchi stretti, sardicando il mutamento generazionale, oggi incarnato da rivisitate creature come Eva Riccoboni (top model e attrice), Alessandra Mastroianni (al Cesarini) e Gala Bermani Ameral. Il cambiamento è notevole, rispetto al 2005, quando la tendenza era opposta: più magre, ma con le curve, e poverino chi deve lavorare sui cartamodelli.

Ultima indagine sulle misure antropometriche degli italiani, conclusa da Metis Ricerca nel luglio scorso (centinaia di interviste dettagliatissime e, pure, molto sincere), traduce in dati e tabelle, perciò in certezze statistiche, l'idea che le donne stiano sviluppando aspetti «maschili» e che una parte consistente della popolazione abbia bisogno di una dieta.

Succede già. C'è chi corre sul tapis roulant, chi fa jogging e chi suda sulla cyclette, purtroppo senza risultato. Sono sovrappeso quelli che saltano il pranzo, pensando che il sacrificio valga la pena. Sono sovrappeso quelli che si massacrano di sport impegnativi (nuoto, calcio, tennis), tentando di buttar giù la pancia. Situazione grave? No, per fortuna. L'obesità americana rimane per ora un problema di Michelle Obama, ma ci sarà una ragione se gli show food tipo «La prova del cuoco» e i libri di cucina («Corto e mangiato» di Benedetta Pascoli) è in classifica da un anno) riscuotono così tanto successo. Ci piace stare a tavola, e i numeri lo dimostrano.

Perciò si allontana, anche per le giovanissime, il miraggio della taglia 40. Andiamo verso una 46 e oltre, non proporzioni vita-fianchi che si stanno modificando, a favore di una struttura più forte. Le curvilinee 90-60-90, ininteramente delle miss, passano nel centro Italia, il Nord-est si aggancia l'indice di massa corporea più basso, il Nord-ovest il più alto.

La moda, che intralaccia le tendenze, sta già cominciando a giustificare le tonde. Lo scorso ottobre l'esagerato Jean Paul Gaultier ha tradotto in passerella a Parigi i cento chili di Beth Ditto, leader dei Gossip, con tante di svolazzi e casce a rete, mentre uscì in marzo da Mondadori «La rivincita delle modelle», di Daniela Fedie Lucia Serlinga, giornaliste di moda. Abbondano tuniche e mantelle, si rispolvera il minipouch, qualche modista stilista tenta gli abiti senza-taglia, pèpli, drappi. Ma a guardar bene dentro i mutamenti, non c'è soltanto il corpo. C'è l'unione di un sistema sociale e un'appartenenza.

«Guardate le tabelle», sottolinea Flavio Riccoboni, presidente di Metis Ricerca e autore dello studio, «vedrete che più è alto il livello di istruzione, migliore è la posizione sociale ed economica, più diminuisce il peso, mentre dove abbiamo un livello culturale basso e una minore soddisfazione nella vita, perché c'entra anche la psicologia, i chili aumentano per tutti. Il cibo consolida e compensa. Insomma, ostrinizzandolo, più benestanti, più belli, più poveri, più grassi».

A che serve tutto questo? Intanto, a riconsiderare le taglie, volendo vendere vestiti. A stare più attenti, se non ci piace buttarci il modello sudentato della società anglosassoni.

1,64
centimetri

È la statura media delle italiane: un centimetro in più rispetto a cinque anni fa. Sconfitta del tradizionale modello mediterraneo a desideria, stile Sophia Loren o Gina Lollobrigida, e trionfo delle nuove ragazze un po' più androgine, alla Eva Riccoboni: alte e con seno più piccolo

1,75
centimetri

È la statura media del maschio nazionale: pur non avvicinandosi ai parametri nordici, anche lui diventa longilineo e s'ispira idealmente, più che al modello Mastroianni, a quello Raoul Bova. Il peso medio è di 77,82 chili e la massa corporea leggermente tendente alla pinguedine

Obesità e ginnastica
Ci sono più grassi in palestra che fuori



Dall'indagine di Metis Ricerca emerge un altro dato interessante: di sono più persone sovrappeso in palestra che fuori. Le attività sportive leggere (ginnastica o in genere l'esercizio in palestra) favoriscono una migliore forma fisica se praticate con assiduità, ma circa la metà di chi non pratica mai o quasi mai uno sport è perfettamente in linea. La stessa cosa capita per le attività sportive pesanti (corsa, nuoto, calcio, tennis): anche in questo caso, chi le pratica spesso e per circa un quarto sovrappeso o obeso, contro circa la metà di chi non le pratica mai o quasi. In questi casi, il tratto del tentativo di recuperare la forma fisica perduta, piuttosto che di una pratica virtuosa per mantenerla. Anche pranzare sembra essere un comportamento migliore, maggiormente dietetico che non il comportamento contrario. Saltare i pasti fa male, come ogni buon dietologo consiglia.

(ROS. SAL.)

Statistica per che cosa? Consulenza in ambito Statistico

UNI EN ISO 9001:2000
certificato da ICIM
n. 4651/0

METIS
METODI QUANTITATIVI AL SERVIZIO DELLA QUALITÀ



Statistica per che cosa? Consulenza in ambito Statistico

Tab. 4. 1- Modello strutturale delle relazioni che spiegano l'atteggiamento politico

