Statistica ufficiale, patrimonio di tutti

Giovanni Barbieri (Istat)

Roma, 20 aprile 2004

1. L'illusione della certezza

Nella vita quotidiana, non c'è situazione in cui non abbiamo bisogno di informazione per prendere una decisione.

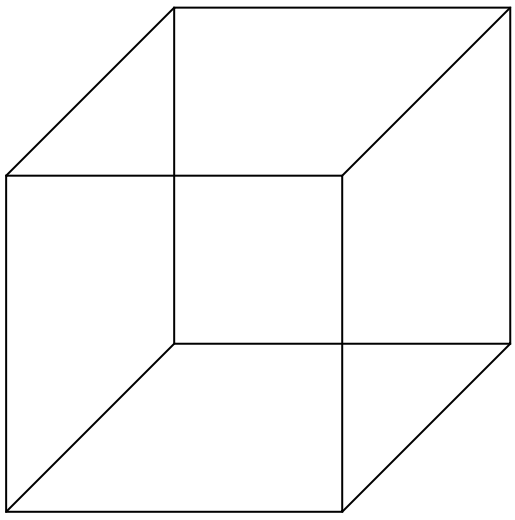
Il più delle volte, non ce ne accorgiamo neppure, perché il nostro sistema nervoso inserisce, per così dire, il pilota automatico e il processo informazione-decisione-azione si svolge al di sotto della nostra soglia di attenzione. È il caso dei riflessi: se qualcuno lancia un oggetto nella nostra direzione, la reazione di abbassarci, incassare la testa nelle spalle e alzare le mani per afferrare o respingere l'oggetto è automatica. Dobbiamo fare uno sforzo di volontà per restare fermi, e percepiamo confusamente lo sforzo mentale che ci è richiesto per vincere quella che chiamiamo una "reazione istintiva". Soggettivamente, ci sembra che nella nostra testa si sia svolto un conflitto tra un "omino" che controlla la nostra mente e la nostra volontà, e il nostro corpo, una specie di animale dominato dagli istinti.

Altre volte, il processo è più consapevole, come nel gesto di afferrare una maniglia, abbassarla e spingere per aprire una porta ed entrare in una stanza. In casi come questi, possiamo analizzare più agevolmente il processo: il nostro sistema percettivo ci informa che siamo di fronte a una porta chiusa; la nostra volontà è quella di entrare nella stanza al di là della porta; le nostre conoscenze (ancora informazione!) comunicano al nostro sistema motorio la sequenza di azioni per aprire la porta. È una descrizione volutamente semplificata e ingenua, ma funzionale al discorso che vogliamo aprire.

© CIRDIS 2004 La riproduzione e l'uso dei materiali sono permessi solo per scopi didattici e non commerciali citando la fonte, gli autori ed i collaboratori.

Ci sono infatti altre situazioni di scelta, meno numerose di quelle evocate finora, in cui le nostre conoscenze e le informazioni che riceviamo sono più scarse e in cui l'esito della decisione è incerto. Sono queste le situazioni di scelta in cui l'informazione offerta dalla statistica è importante.

Prima, però, vorrei utilizzare gli esempi che ho fatto per sfatare un'illusione: l'illusione della certezza. Nella maggior parte dei casi, che sono (quantitativamente, non per importanza percepita) del tipo di quelli descritti in precedenza, ci comportiamo come se i dati della nostra percezione fossero certi, cioè non soggetti ad alcuna ambiguità. Le neuroscienze ci hanno spiegato da tempo che non è così: non c'è nessun omino seduto nella nostra glandola pineale – come pensava Cartesio – che guarda su uno schermo le immagini inviate dalla retina. Al contrario, il nostro sistema nervoso, nell'interazione di numerosi sotto-sistemi, riceve i segnali che provengono dagli occhi, li analizza sotto i profili del movimento, del colore, del contrasto tra bordi e sfondo, li integra con "modelli" e rappresentazioni preesistenti conservate in memoria e crea l'illusione della certezza.



Molti esempi possono illustrare la nostra tendenza a creare certezza, ma qui basterà un'illusione ottica famosa e molto semplice: il cubo di Necker (creato nel 1832 dal cristallografo svizzero Louis Albert Necker).

La profondità del cubo è ambigua, perché l'immagine è una rappresentazione bidimensionale e mancano quegli artifici grafici (linee nascoste, ombreggiature) che in genere ci permettono di dire quale faccia è davanti e quale dietro. Ma noi non percepiamo alcuna ambiguità: vediamo la faccia più a sinistra come anteriore e l'altra come posteriore, o viceversa; ma mai contemporaneamente. Dopo un po' si verifica un fenomeno detto "Gestalt switch" e vediamo davanti quello che prima vedevamo dietro, come se il cubo girasse nello spazio. Mai contemporaneamente.

Un po' inquietante? Ma, a rifletterci, meglio così. Nella vita di tutti i giorni, l'illusione della certezza ci aiuta a prendere migliaia di micro-decisioni rapidamente e senza mettere a repentaglio la nostra sicurezza. Ma è importante, per quello che diremo dopo, avere acquisito la consapevolezza che siamo dotati di un "meccanismo inferenziale" che analizza informazioni incomplete e ambigue e ne estrae la "migliore stima" come sostegno alle decisioni.

Affrontiamo adesso un altro contesto di decisioni, un po' meno elementari ma altrettanto familiari: quale auto mi conviene comprare? Dove investire i miei risparmi? Quale facoltà universitaria mi dà sbocchi professionali interessanti? Mi faccio operare agli occhi? Domani pioverà? Vado a vivere in campagna? La patente a punti ha fatto diminuire gli incidenti stradali? È vero che chi si fa le canne poi passa all'eroina? Dopo l'11 settembre, è più sicuro viaggiare in macchina o in aereo? Chi ci amministra ha mantenuto i suoi impegni?

In questi esempi è molto più evidente che l'esito della decisione è incerto e che l'informazione a disposizione insufficiente, ma anche che, nell'impossibilità di conoscere con certezza l'esito della nostra decisione, è comunque meglio poter valutare la probabilità degli esiti possibili, sulla base di un gran numero di situazioni simili.

Questa è la radice della statistica come scienza.

2. Che cos'è la statistica

Secondo il Vocabolario Treccani, la statistica è la

> scienza che ha per oggetto lo studio dei fenomeni collettivi suscettibili di misura e di descrizione quantitativa (specialmente quando il numero degli individui interessato è talmente elevato da escludere la possibilità o la convenienza di seguire le vicende di ogni singolo individuo): ba-

> sandosi sulla raccolta di un grande numero di dati inerenti ai fenomeni in esame, e partendo da ipotesi più o meno direttamente suggerite dall'esperienza o da analogie con altri fenomeni già noti, mediante l'applicazione di metodi matematici fondati sul calcolo delle probabilità, si perviene alla formulazione di leggi di media che governano tali fenomeni, dette *leggi statistiche*; spesso la raccolta dei dati viene limitata a un *campione* più ristretto, opportunamente predeterminato in modo da rappresentare fedelmente le caratteristiche generali.
> Concepita inizialmente come attività descrittiva di certi fatti sociali e in particolare come attività amministrativa dello Stato, ha via via ampliato i suoi confini, fino a diventare una vera e propria "scienza del collettivo", disciplina con finalità non solo descrittive dei fenomeni sociali e naturali, ma orientata anche a finalità di ricerca; si parla così di *statistica descrittiva*, che si riferisce alla classificazione ed elaborazione dei dati, e di *statistica inferenziale*, che riguarda l'interpretazione dei dati e che permette di trarre indicazioni sui fenomeni non osservati direttamente. […]

Il primo punto da sottolineare in questa definizione è il rapporto tra "statistica" e "fenomeni collettivi": il tipo di incertezza che la statistica cerca di tenere sotto controllo è insita nel numero degli individui o dei fenomeni.

Il secondo è che la strategia per ridurre questa fonte di incertezza è quella di semplificare la complessità attraverso un modello – basato sul calcolo delle probabilità – capace di conservare gran parte dell'informazione riducendo drasticamente la molteplicità dei fenomeni.

Torneremo su questo aspetto perché spesso si dimenticano i rapporti strettissimi della statistica con la rivoluzione dell'informazione che stiamo vivendo in questi anni e con la società dell'informazione che sta nascendo.

Il terzo aspetto da sottolineare è il riferimento alla misurabilità: se la statistica non fosse in grado di descrivere quantitativamente la realtà, non potrebbe fregiarsi del titolo di scienza. A questo fine, è importante introdurre una distinzione tra "incertezza" e "rischio".

Si usa definire "incerta" una situazione nella quale, per mancanza di evidenza empirica, non è possibile assegnare un valore di probabilità ai possibili risultati alternativi. L'incertezza si trasforma in "rischio" se invece, sulla base di una qualche evidenza empirica, è possibile assegnare un valore di probabilità o di frequenza agli esiti possibili. Non è necessario che questo valore sia dato una volta per tutte; può cambiare o diventare più preciso alla luce di nuova informazione e nuova esperienza.

Esistono (almeno) tre versioni del concetto di probabilità:

- la *frequenza* relativa di un evento all'interno di una certa classe di riferimento, calcolata sulla base di un grande numero di osservazioni: ad esempio, la frequenza relativa del cancro ai polmoni nei maschi italiani che hanno fumato almeno un pacchetto di sigarette al giorno per almeno vent'anni;

- la *propensione*, cioè la probabilità legata alle caratteristiche fisiche di un oggetto: in una moneta a due facce la probabilità che esca testa è 0,5; in un dado a sei facce la probabilità che esca tre è circa 0,17;

- la *probabilità soggettiva*, in cui la probabilità di un evento è la misura del *grado di fiducia* che un individuo coerente attribuisce, secondo le sue informazioni e opinioni, al suo avverarsi.

Senza entrare nelle dispute che – come immaginerete – dividono gli addetti ai lavori, mi limito a due osservazioni: che il concetto di rischio si applica più chiaramente e univocamente alle prime due definizioni di probabilità, perché probabilità soggettive possono essere assegnate anche a eventi unici o nuovi; ma che, al momento della decisione, il soggetto che deve assumerla ragiona sempre, in ultima istanza, in termini di probabilità soggettiva.

3. Statistica e democrazia: un profilo storico

L'ignoranza pubblica del rischio ha una base storica.

A differenza della narrativa – dalla mitologia alla storia, passando attraverso il romanzo, la poesia e il pettegolezzo – che è elemento essenziale del nostro modo di pensare dai primordi dell'umanità, l'approccio quantitativo ai fenomeni in generale, e la statistica pubblica in particolare, sono acquisizioni culturali relativamente recenti.

Il potere insito nell'informazione statistica è noto da secoli ai governi. Ancora una volta ci viene in soccorso il Vocabolario Treccani, che ci informa anche che il nome è il femminile sostantivato dell'aggettivo "statistico", il cui significato originario era "che riguarda lo Stato, la vita e i problemi dello Stato".

Per tutto il Settecento e gran parte dell'Ottocento, l'informazione statistica era un segreto di Stato, nota solo a un'*elite* e tenuta nascosta al pubbli-

co. L'appetito di Napoleone per i dati del suo *bureau de statistique* è leggendario: alla corte di Napoleone si diceva che per ottenere qualcosa dall'imperatore, era sufficiente dargli in cambio delle statistiche.

La volontà di rendere pubblici i dati economici e demografici è piuttosto recente, e risale più o meno al 1830. La crescente disseminazione di informazione statistica ai cittadini – durante l'Ottocento e il Novecento – è stata associata alla diffusione della democrazia nell'Occidente.

Oggi siamo sommersi da una valanga di numeri. Una delle caratteristiche delle democrazie è di rendere disponibile una grande massa di informazioni statistiche ai cittadini. Inoltre, in tutti i Paesi democratici, la statistica ufficiale è finanziata con i soldi dei contribuenti e messa a disposizione di tutti i cittadini.

Questo *status* della statistica pubblica pone almeno due problemi. Primo: perché la statistica gode di una posizione simile a quella dei servizi pubblici essenziali, anche se a un primo sguardo, la statistica appare meno necessaria dell'assistenza sanitaria o della pubblica sicurezza? Secondo: perché, a fronte dell'onnipresenza dell'informazione statistica ufficiale, i cittadini sono così selettivi nei loro interessi e diffidenti nell'uso delle statistiche rese disponibili?

4. Statistica ed eguaglianza delle opportunità

La risposta al primo quesito è legata alla natura delle decisioni per le quali l'approccio statistico è di ausilio: si tratta di decisioni soggette a rischio, in cui il numero dei fenomeni o degli individui è talmente elevato da escludere la possibilità o la convenienza di seguire le vicende di ogni singolo individuo, e occorre basarsi sulla raccolta di un grande numero di dati inerenti ai fenomeni in esame per poter assegnare un valore di probabilità agli esiti possibili. Il costo della raccolta delle informazioni elementari che sono poi riassunte nelle statistiche e che consentono la quantificazione del rischio è molto elevato, e non tutti possono permetterselo. Questo è il motivo per cui, in origine, la statistica operava esclusivamente al servizio del potere politico. In una società democratica, i cittadini devono essere posti in una posizione di eguaglianza iniziale: l'eguaglianza delle opportunità e delle possibilità. Se così non fosse, i soggetti più deboli (individui, imprese, istituzioni, gruppi sociali) sarebbero in una posizione di svantaggio, non potendo "acquistare" informazione statistica sul mercato, rispetto a quelli più forti. Sussistono, in altri termini, delle "asimmetrie informative". Di conseguenza, lo Stato provvede a produrre statistiche ufficiali

e pubbliche – cioè certificate e messe a disposizione di tutti – come strumento di garanzia dell'eguaglianza dei diritti dei cittadini.

Inoltre, mettere a disposizione di tutti le stesse informazioni di cui dispongono coloro che ci governano è anche un modo per permettere di valutare l'efficienza e l'efficacia dell'azione pubblica.

Per questo, la statistica ufficiale è uno strumento essenziale della democrazia.

5. L'ignoranza dell'incertezza

La risposta alla seconda domanda va probabilmente ricercata in quanto abbiamo detto a proposito dell'illusione della certezza. La propensione a ignorare l'incertezza è innata, mentre essere informati sul rischio è un fenomeno recente e culturalmente determinato.

Nelle società democratiche si osserva una maggiore tendenza a usare informazioni quantitative, ad associarle a contesti decisionali, a rendere i rischi trasparenti.

Ad esempio, negli Stati Uniti viene assegnato un valore numerico alla probabilità di pioggia nelle previsioni meteorologiche già dal 1965, mentre in Germania questa prassi è diventata abituale soltanto dal 1990 e in Italia è ancora piuttosto raro.

Di conseguenza, anche se la statistica ufficiale è resa pubblica, i cittadini spesso non hanno gli strumenti culturali per metterla a buon uso "*Numbers are public, but the public is not generally numerate*", secondo la sintesi di Gigerenzer.

Molte circostanze storiche e molti fenomeni culturali hanno condotto a questa situazione. Innanzitutto, nella società operano gruppi di pressione che non hanno interesse alla diffusione della *numeracy*, cioè dell'alfabetizzazione statistica. Ad esempio, negli anni Cinquanta i produttori di tabacco americani intrapresero una campagna per convincere i consumatori che fumare non provocava danni alla salute, anche se la comunità scientifica stava raccogliendo prove decisive sulla relazione tra fumo, cancro al polmone e malattie cardiovascolari. Si stima che questa campagna in difesa dell'*illusione della certezza* sia costata centinaia di migliaia di dollari. Dopo il rapporto del Surgeon General, pubblicato nel 1964, i produttori di tabacco intrapresero una seconda campagna, questa

volta insinuando il dubbio sull'attendibilità delle conclusioni (statistiche) dei rischi connessi al fumo. Questo esempio è paradigmatico delle strategie messe in campo – spesso con successo – da chi ha interesse a ostacolare la *numeracy* e la consapevolezza informata dei cittadini: primo, rafforzare e se necessario costruire l'illusione della certezza; secondo, quando l'illusione crolla e l'*incertezza* è riconosciuta, insinuare il dubbio sulla possibilità di stimare correttamente i *rischi*. Questa strategia è stata utilizzata molte volte, spesso con successo, ed è una delle ragioni per le quali, in una società saldamente e realmente democratica, è bene che la statistica sia pubblica e sottratta all'influsso dei gruppi di interesse.

Una seconda spiegazione è che l'illusione della certezza è radicata profondamente nel modo di operare della nostra mente e ha vantaggi tali, nella maggior parte delle situazioni, da orientare il nostro modo di pensare. Per questo, tutti troviamo il modo di ragionare probabilistico particolarmente ostico.

Anche in questo caso è utile un esempio.

5.1 *Una digressione: il problema di Monty Hall*

Il *Problema di Monty Hall* riguarda una situazione di gioco classica e prende nome da Monty Hall, conduttore per lunghi anni della trasmissione televisiva americana *Let's Make a Deal*. Si hanno tre porte indicate con numeri da 1 a 3. Dietro una delle porte c'è un'automobile, dietro ciascuna delle altre delle capre.

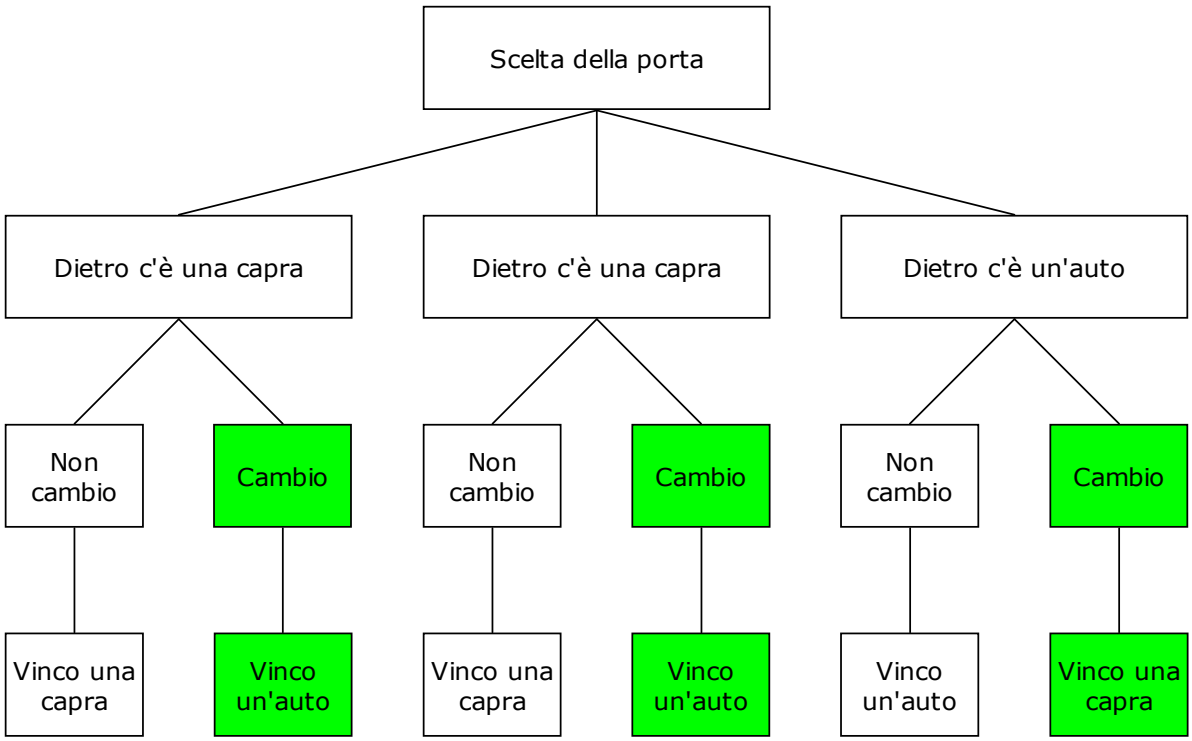
Il concorrente ha la possibilità di scegliere una delle porte e gli sarà assegnato come premio quello che è nascosto dietro la porta prescelta.

Il gioco è diventato un celebre e controverso problema di calcolo delle probabilità a seguito di alcuni articoli di Marilyn Vos Savant (detentrice di un primato sul Guinness per il suo quoziente d'intelligenza) nella rubrica *Ask Marilyn* del *Parade magazine*, un popolare supplemento domenicale. La controversia ebbe inizio quando nel settembre del 1990 un lettore, Craig F. Whitaker, pose il problema nei seguenti termini:

> supponiamo che – dopo la scelta del concorrente – il presentatore, che sa dove sono l'automobile e le capre, apra una delle altre due porte e mostri che dietro c'è una capra. A questo punto al concorrente viene offerta la possibilità di confermare la sua scelta o cambiarla: quale è la sua strategia migliore, se il suo obiettivo è quello di vincere la macchina? Confermare la prima scelta o cambiare?

Marilyn Vos Savant rispose che conveniva sempre cambiare, perché in quel caso le probabilità di vincere l'auto salivano al 67 per cento (due su tre). Decine di persone, tra cui molti matematici, scrissero alla rivista: il 92 per cento sostenendo che la risposta era sbagliata.

In effetti ha ragione Marylin Vos Savant, e lo si può dimostrare matematicamente e anche – secondo me in modo molto più elegante – con un grafico dell'*albero delle decisioni*.



Perché risulta così difficile, anche per persone colte e istruite, e persino per i matematici, trarre inferenze corrette sulla base del calcolo delle probabilità? Secondo molti, dipende dal fatto che la teoria della probabilità, che ha per oggetto i processi di inferenza da informazione incompleta e incerta, è acquisizione relativamente recente. Per chi ama le date, la si usa far nascere nel 1654, anno di un famoso scambio di lettere tra il filosofo (e matematico) Blaise Pascal e il matematico Pierre Fermat (quello del famoso "ultimo teorema") sul gioco d'azzardo. Ian Hacking ha definito la scoperta tardiva della probabilità, dopo che la maggior parte degli altri strumenti chiave del conoscere erano stati analizzati e formalizzati, uno scandalo della filosofia. La difficoltà che persino grandi pensatori fronteggiavano prima delle scoperte di Pascal e Fermat è ben illustrato da *De Ludo Aleae* di Girolamo Cardano (scritto tra il 1525 e il 1565): appassiona-

to giocatore (agli inizi della carriera si manteneva grazie alle vincite al gioco), Cardano introdusse il termine "probabilità" (che contiene nella sua radice il riferimento al concetto di "prova"), ma sostenne – sulla base delle caratteristiche fisiche del dado, e dunque all'interno di una concezione del *rischio* basata sulla *propensione* – che ogni faccia di un dado a sei facce sarebbe risultata esattamente una volta in sei lanci. Di fronte a risultati empirici evidentemente diversi, risolse il problema introducendo il concetto di fortuna!

Personalmente, sospetto che le cause siano ancora più profonde, e legate al funzionamento innato del pensiero umano. Tutti noi possediamo un modello innato dei fenomeni fisici ordinari, grazie al quale calcoliamo immediatamente e senza apparente sforzo la traiettoria di una palla che rimbalza, ma non un modello innato della fisica quantistica (che ci appare assolutamente contro-intuitiva) né delle regole della probabilità (che ci appaiono altrettanto ostiche).

Siamo quindi arrivati a una conclusione importante. Date le premesse che abbiamo appena illustrato, il ruolo della statistica pubblica non è soltanto quello di mettere a disposizione dei cittadini le informazioni necessarie ad assumere decisioni consapevoli, ma anche quello di comunicare queste informazioni in modo chiaro e non fuorviante, diffondendo i necessari strumenti di interpretazione e di conoscenza e operando per radicare la *statistical literacy* nella società dell'informazione.

6. La statistica pubblica e i suoi principi

Che cos'è, dunque, la statistica pubblica? E quali sono i suoi principi?

La statistica pubblica offre ai cittadini un'immagine non distorta della società e dell'economia; è utile per discutere le alternative e per assumere le decisioni; offre la possibilità di conoscere e valutare le attività e la *performance* degli amministratori.

È importante sottolineare che i concetti di statistica ufficiale e di statistica pubblica, ancorché strettamente correlati, non coincidono:

- il concetto di *statistica ufficiale* fa riferimento ai soggetti che la producono e diffondono, e in ultima istanza al mandato istituzionale di cui essi sono investiti dalla legge (in Italia, è statistica ufficiale quella prevista ogni anno dal Programma statistico nazionale);

- il concetto di *statistica pubblica* fa riferimento alle persone che forniscono il servizio, che sono dipendenti pubblici (*civil servants*, come si dice nella tradizione anglosassone); alle sue attività (raccogliere i dati, elaborarli, diffonderli), che sono al servizio dei cittadini; ai suoi prodotti e servizi, che hanno la natura di bene pubblico.

I principi fondamentali della statistica ufficiale sono stati adottati dapprima dalla Commissione economica per l'Europa delle Nazioni unite con la Decisione C47 del 15 aprile 1992 e successivamente fatti propri (con un nuovo preambolo) dalla Commissione statistica dell'ONU nella sessione plenaria dell'11-14 aprile 1994.

> I. Le statistiche ufficiali costituiscono un elemento indispensabile nel sistema informativo di una società democratica. Esse sono poste al servizio delle istituzioni, degli operatori economici e del pubblico ai quali forniscono dati sulla situazione economica, demografica, sociale e ambientale. A questo fine, le statistiche ufficiali che risultano di utilità generale debbono essere elaborate e rese disponibili in modo imparziale dalle istituzioni della statistica ufficiale affinché sia soddisfatto il diritto dei cittadini all'informazione pubblica.

In particolare, imparzialità implica che la statistica pubblica deve essere prodotta in modo oggettivo e indipendente; deve essere libera da pressioni di partiti politici e altri gruppi d'interesse; che la scelta delle tecniche, delle definizioni e delle metodologie deve essere orientata soltanto dagli obiettivi perseguiti; e che, infine, le statistiche prodotte devono essere rese disponibili – con la massima tempestività possibile – a tutti gli utenti (Governo, altre amministrazioni, operatori economici e sociali, università ed enti di ricerca, cittadini).

> II. Allo scopo di mantenere la fiducia da parte degli utenti nelle statistiche ufficiali, le istituzioni statistiche debbono assumere le loro decisioni sulla base di considerazioni strettamente professionali, ivi compresi i principi scientifici e di etica professionale, sui metodi e sulle procedure per la raccolta, l'elaborazione, l'immagazzinamento e la presentazione dei dati statistici.

La fiducia dei cittadini nell'affidabilità e nell'imparzialità delle statistiche ufficiali è cruciale. Non è un caso che *Statistics: a Matter of Trust* sia il titolo che il Governo del Regno Unito ha dato al Libro verde con il quale apriva la discussione sulla rifondazione del suo sistema statistico nazionale:

- l'affidabilità della statistica ufficiale è una pietra angolare della democrazia ed è essenziale per la corretta gestione della cosa pubblica, per la sua trasparenza e per l'individuazione delle responsabilità;

- i cittadini, il Governo, il Parlamento e l'insieme delle pubbliche amministrazioni hanno tutti interessi legittimi in statistiche affidabili, che offrano una rappresentazione obiettiva dell'economia e della società, che siano comparabili nel tempo e che forniscano un adeguato dettaglio territoriale;

- le statistiche forniscono al Governo e alle pubbliche amministrazioni uno strumento per la formulazione e la valutazione delle politiche e per la gestione dei servizi e delle funzioni pubbliche;

- d'altro canto, le statistiche forniscono ai cittadini uno strumento di valutazione sull'azione dell'amministrazione pubblica e sui risultati da essa conseguiti;

- salvaguardare l'affidabilità delle statistiche pubbliche e la fiducia dei cittadini è una responsabilità di primaria importanza.

III. Per facilitare una corretta interpretazione dei dati, le istituzioni statistiche debbono presentare l'informazione rispettando gli standard scientifici relativi alle fonti, ai metodi e alle procedure della statistica.

I concetti, i metodi e le regole relativi alla rilevazione, all'elaborazione e alla diffusione d'informazioni statistiche sono fissati in base alle norme professionali, ai metodi scientifici e alla deontologia, in modo che i risultati statistici rappresentino fedelmente la realtà.

IV. Alle istituzioni statistiche deve essere riconosciuto il diritto di fornire commenti sulle interpretazioni erronee e sulle utilizzazioni non corrette.

Questo articolo implica un duplice diritto-dovere di responsabilità. Primo, gli statistici pubblici sono tenuti a opporsi a qualsiasi tipo di rilevazione, elaborazione, analisi e presentazione di dati che potrebbe dar adito a interpretazioni errate. Secondo, gli organi della statistica pubblica sono autorizzati a commentare interpretazioni fuorvianti e utilizzazioni abusive dei loro risultati statistici.

V. I dati raccolti a fini statistici possono derivare da ogni tipo di fonte, sia essa una indagine statistica o una documentazione amministrativa. Le istituzioni statistiche debbono poter scegliere la fonte più appropriata in relazione alla qualità, alla tempestività, ai costi e agli oneri sui rispondenti.

Attualmente, questo articolo viene interpretato in modo stringente: le rilevazioni sono realizzate solo se i dati amministrativi non sono disponibili o sono insufficienti. Inoltre, le rilevazioni sono condotte in modo da ridurre al minimo l'onere dei rispondenti.

Va inoltre garantita l'economicità della statistica pubblica, che consiste nell'uso ottimale delle risorse disponibili: i costi della produzione statistica pubblica (in termini finanziari e di uso delle risorse umane) devono pertanto essere valutati in termini di costi e di benefici.

VI. I dati individuali raccolti dalle istituzioni statistiche per la produzione di statistiche ufficiali, sia che si riferiscano a persone fisiche che a persone giuridiche, debbono restare strettamente confidenziali e debbono essere usati esclusivamente per scopi statistici.

Questo articolo fa riferimento a tre punti importanti: quello del *segreto statistico* (gli organi della statistica pubblica trattano in modo strettamente confidenziale i dati su singole persone fisiche o giuridiche; non divulgano alcuna informazione che consenta di risalire alla situazione di una persona fisica o giuridica), quello dell'*utilizzazione esclusiva* (i dati sulle persone fisiche o giuridiche rilevati a fini statistici devono essere usati esclusivamente per scopi statistici e non possono essere utilizzati per decisioni o misure amministrative concernenti queste persone), quello dell'*informazione* (i rispondenti sono informati in merito alle basi e agli obiettivi delle rilevazioni, nonché alle misure adottate in materia di protezione dei dati).

Il riferimento alla riservatezza implica, in particolare, che i dati riferiti a una singola unità statistica – ottenuti direttamente da una rilevazione o indirettamente da fonti amministrative – devono essere protetti da ogni violazione del diritto alla riservatezza e che devono essere adottate misure di prevenzione attiva rispetto all'uso non statistico dei dati.

VII. Le leggi, i regolamenti e i metodi di misura in base ai quali operano i sistemi statistici debbono essere resi pubblici.

La raccolta di dati su persone fisiche o giuridiche poggia su una base legale.

La produzione statistica pubblica è guidata da obiettivi e programmi chiaramente definiti (programma statistico nazionale, programma statistico dell'Unione europea). I programmi determinano le aree di ricerca, il calendario e l'ampiezza delle rilevazioni. Attenzione particolare è dedicata ai nuovi sviluppi demografici, economici, sociali e ambientali.

Oltre che alla base giuridica, questo articolo fa dunque implicitamente riferimento anche al principio della *pertinenza*.

La statistica ufficiale infatti mira a riprodurre la realtà che intende rappresentare nel modo più fedele possibile. Per questo, la selezione di fonti, metodi e procedure è guidata da criteri scientifici e le informazioni sulla copertura, la metodologia, le procedure e le fonti adottate contribuiscono a migliorare l'affidabilità delle statistiche.

Le modalità con cui il sistema statistico pubblico opera sono di pubblico dominio, perché soltanto gli utilizzatori sono in grado, nel loro complesso, di valutare la pertinenza delle informazioni prodotte.

VIII. Il coordinamento tra le istituzioni statistiche nell'ambito di un Paese è essenziale per ottenere coerenza ed efficienza nel sistema statistico.

IX. Nell'ambito di ogni paese l'utilizzo da parte delle istituzioni statistiche di concetti, classificazioni e metodi definiti a livello internazionale promuove la coerenza e l'efficienza dei sistemi statistici a tutti i livelli ufficiali.

X. La cooperazione bilaterale e multilaterale in materia statistica contribuisce allo sviluppo dei sistemi della statistica ufficiale in ogni Paese.

La compatibilità delle informazioni è assicurata all'interno e tra i diversi settori della statistica pubblica. A tal fine, si applicano i concetti, le classificazioni, le terminologie e i metodi stabiliti e riconosciuti a livello nazionale o internazionale.

Nell'ambito di ogni paese, l'uso da parte delle istituzioni statistiche di concetti, classificazioni e metodi definiti a livello internazionale promuove la coerenza e l'efficienza dei sistemi statistici a tutti i livelli.

7. Statistica pubblica e società dell'informazione

Per concludere, vorrei tornare – come avevo promesso – ai rapporti tra statistica pubblica e società dell'informazione.

Un primo punto che vorrei sottolineare è che a mio parere l'aspetto più stimolante della rivoluzione dell'informazione non sono gli elementi tecnologici, ma le possibilità di costruire una società più aperta.

L'informazione, oggi, è la cosa più simile al pensiero che sia mai esistita: i suoi prodotti si avviano a diventare indipendenti dal loro supporto fisico, a essere conservati e trasmessi non come un oggetto, ma come una serie di zero e di uno, come un flusso di variazioni di voltaggio che si sposta alla velocità della luce, in cui la distinzione tra originale e copia cessa di avere significato.

L'informazione digitale è (potenzialmente e sempre più) conservata su un meta-supporto che ha la medesima natura dell'informazione stessa. Gli oggetti fisici su cui l'informazione digitale risiede non hanno valore né rilevanza in quanto tali; l'utente è interessato alla manifestazione dell'informazione, ma non ne entra in possesso nel senso tradizionale del termine.

In questo modo, le caratteristiche specifiche dell'informazione sembrano sfuggire a molte delle categorie che si applicano alle merci, e su cui si esercitano le leggi dell'economia e le regole che ne governano l'ordinato funzionamento.

In primo luogo, l'informazione si comporta come una forma di vita (o forse è la vita a essere una forma di informazione). Il biologo inglese Richard Dawkins ha proposto venticinque anni fa l'idea di *meme*, l'unità minima di informazione capace di duplicarsi e di propagarsi negli ecosistemi della memoria. I memi, come i geni, si riproducono, interagiscono con l'ambiente, si adattano, occupano nicchie, persistono. L'informazione sperimenta continue mutazioni, anche se immagazzinarla su un supporto statico ostacola questo processo: la trasmissione orale e quella digitale dell'informazione, a due estremi opposti, favoriscono l'evoluzione continua delle informazioni, senza che possa esistere una loro edizione definitiva. D'altra parte, proprio per questo l'informazione tende a essere altamente deperibile, anche se la durata della sua vita utile dipende in larga misura dalla sua natura e dagli scopi dei suoi fruitori.

In secondo luogo, l'informazione è meglio definita come un'attività, piuttosto che come un oggetto. Liberata dal supporto che la veicola, si manifesta come un'interazione tra menti umane. Nella definizione di Gregory Bateson, "*information is a difference which makes a difference*": questa differenza è un'attività che interviene nell'ambito di una relazione, è un'azione che accade nel tempo piuttosto che un oggetto che occupa uno spazio. L'informazione può essere immagazzinata in un supporto fisico, ma esiste soltanto in una transazione: per manifestarsi deve essere "decompressa" dal supporto in cui è codificata e sperimentata da chi ne è destinatario. Inoltre, l'informazione partecipa della natura dei beni pubblici ("non rivalità" nella fruizione e "non escludibilità" dei fruitori); non si trasferi-

sce, ma si propaga, lasciando una traccia dietro di sé: se vendo un'informazione, entrambi la conosciamo.

In terzo luogo, l'informazione si manifesta entro un rapporto sociale. Il valore dell'informazione risiede nel suo significato, nel luogo e nel momento in cui la trasmissione diviene ricezione, nell'interazione tra due attori sociali. Il valore di ciò che è trasmesso dipende interamente dalla disponibilità di recettori adeguati da parte del fruitore: attenzione, interesse, linguaggio, terminologia, paradigmi condivisi. Lo scarto tra ricezione e comprensione, cioè il significato, è la differenza tra dati e informazione. In questo contesto, la consueta relazione inversa tra scarsità e valore, così importante per i beni materiali, non sussiste più; al contrario, in molti casi sono la familiarità e la reputazione a dare più valore a un'informazione. Operano in questa direzione due meccanismi. Il primo ha a che fare con la creazione di standard di fatto: più persone comunicano utilizzando il medesimo sistema di convenzioni e di informazioni (un linguaggio, un paradigma scientifico, una base di dati, un sistema di valori), più questo sistema assume valore in quanto condiviso da una platea più vasta di soggetti che comunicano. Il valore di una rete cresce in ragione del numero di interconnessioni possibili, e non del numero di nodi. Il secondo ha a che fare con l'autorevolezza della fonte dell'informazione: in una situazione in cui l'informazione è sovrabbondante, viene attribuito valore maggiore all'informazione proveniente da una fonte che in passato ha dimostrato di essere attendibile, di fornire informazione pertinente e di buona qualità. L'informazione, come la statistica, è una questione di fiducia.

Tutti questi elementi convergono a orientare verso una conclusione in qualche misura scontata, che però trae spessore dalle considerazioni svolte: l'informazione va intesa tanto dai produttori quanto dai fruitori come un servizio, il cui valore si fonda sulla possibilità di un accesso interattivo del fruitore alla competenza del produttore.

Il riferimento a una concezione dell'informazione come servizio contribuisce a chiarire le condizioni del passaggio dalle tecnologie dell'informazione alla società della conoscenza. È l'utilità dell'informazione percepita dall'utente, la sua pertinenza commisurata ai bisogni conoscitivi, la possibilità di utilizzare l'informazione per assumere decisioni riducendo l'incertezza e la complessità a rappresentare il primo requisito della trasformazione dell'informazione in conoscenza. Semplificando questo passaggio in uno schema concettuale, la transizione è dal dato (la rappresentazione e la misura di un fenomeno colte nel momento della loro produzione), all'informazione (che si manifesta nella comprensione/interpretazione del dato, cioè nell'interazione tra soggetti sociali e nell'attribuzione di signifi-

cato da parte del destinatario), alla conoscenza (cioè all'attribuzione di valore all'informazione, in funzione del soddisfacimento di un bisogno, anche di conoscenza in sé, e dell'utilizzazione nell'assunzione di una decisione).

In realtà, il processo conoscitivo è più complesso. L'informazione, per trasformarsi in conoscenza, deve largamente essere interiorizzata, a livello tanto degli individui, quanto dei soggetti sociali. Deve potere interagire con altri elementi informativi, in modo da costituire un reticolo di informazione interconnessa. Deve potere essere utilizzata all'interno di processi che creano "metafore", cioè meccanismi atti ad applicare l'informazione o complessi di informazioni a problemi e ad ambiti di applicazione diversi rispetto a quelli per i quali l'informazione era stata originariamente raccolta. Questi processi, che sono quelli attraverso i quali la conoscenza si produce a livello individuale, sono anche quelli che governano i processi sociali di creazione di conoscenze e di competenze, all'interno dei processi produttivi, nelle comunità locali, negli stili di vita, nella creazione di universi culturali, nel governo della cosa pubblica.

Questa complessità, riconducendoci alla fine al carattere sociale e sistemico dei processi conoscitivi, pone una serie di problemi che investono chi ha la responsabilità, in quanto operatore pubblico, di alimentare la conoscenza e di vegliare affinché a tutti i cittadini sia consentito un accesso adeguato all'informazione.

Dal punto di vista della statistica ufficiale, la riduzione della complessità e la ricerca di *pattern* dotati di significato sono un compito specifico del lavoro statistico; ma il rischio che, a fronte di giacimenti informativi tanto vasti, questi processi possano condurre a una selezione dell'informazione capace di distorcerla è reale, e trascende il terreno della correttezza metodologica per investire quello della deontologia. Il pericolo più grave è quello che le semplificazioni operate dalla statistica ufficiale possano essere percepite non come un servizio pubblico inteso alla documentazione e all'analisi dei fenomeni, ma come una manipolazione della realtà stessa guidata da valutazioni soggettive o, peggio, interessate. Verrebbe meno, in tal caso, quella fiducia che è alla base del patto sociale su cui si fonda la statistica ufficiale.

È altrettanto necessaria una crescita delle capacità di orientarsi dell'utenza, su cui graverà inevitabilmente una parte crescente degli oneri di vigilanza e di controllo sull'informazione disponibile in rete. Gli utenti dovranno essere in grado di discernere l'autorevolezza delle fonti anche al di là di quelle certificate, perché altrimenti si ostacolerebbero la scoperta e lo sviluppo di nuovi giacimenti informativi; dovranno sviluppare le loro

capacità di valutazione dei limiti di validità di determinate informazioni; dovranno saper orientarsi e discriminare tra informazioni in conflitto, magari soltanto apparente; dovranno maturare un'autonoma capacità interpretativa. In questo processo, non dovranno essere lasciati soli dalla statistica ufficiale: la società dell'informazione richiede un grande sviluppo della cultura statistica, e più in generale della *numeracy*, e dunque un investimento formativo su cui la scuola e la statistica ufficiale dovranno cooperare strettamente.

Riferimenti bibliografici

BARLOW, John Perry (1994). "The Economy of Ideas". *Wired* 2.03.

BARLOW, John Perry (2000). "The New Economy of Ideas". *Wired* 8.10.

BATESON, Gregory (1972). "Form, Substance, and Difference". In: *Steps to an Ecology of Mind*. New York. Ballantine Books.

BLACKMORE, Susan (1999). *The Meme Machine*. Oxford. Oxford University Press.

DAWKINS, Richard (1976). *The Selfish Gene*. Oxford. Oxford University Press.

DYSON, Esther (1994). "Intellectual Property on the Net". *Release 1.0* (December 1994).

DYSON, Esther (1997). *Release 2.0: A Design for Living in the Digital Age*. New York. Broadway Books.

ECONOMIC SECRETARY TO THE TREASURY (1998). *Statistics, a Matter of Trust. A Consultation Document*. London. The Stationery Office

GIGERENZER, Gerd (2003). *Calculated Risks: How to Know When Numbers Deceive You*. New York. Simon & Schuster.

HACKING, Ian (1984). *The Emergence of Probability: A Philosophical Study of Early Ideas about Probability, Induction and Statistical Inference*. Cambridge. Cambridge University Press.

HACKING, Ian (1990). *The Taming of Chance*. Cambridge. Cambridge University Press.

HADDON, Mark (2003). *The Curious Incident of the Dog in the Night-time*. London. Jonathan Cape.

LEWIS, Harold Warren (1998). *Why Flip a Coin?: The Art and Science of Good Decisions*. New York. John Wiley & Sons.

PINKER, Steven (1999). *How the Mind Works*. Harmondsworth. Penguin.

PORTER, Theodore M. (1988). *The Rise of Statistical Thinking, 1820-1900*. Princeton. Priceton University Press.

PORTER, Theodore M. (1996). *Trust in Numbers*. Princeton. Priceton University Press.

STIGLITZ, Joseph (1988). *Economics of the Public Sector* (2nd ed.). New York. Norton.

UNITED NATIONS-ECONOMIC COMMISSION FOR EUROPE (1992). *Principles of Official Statistics*. Geneva.