

BIBLIOTECA
DOCUMENTAZIONE
RELAZIONI INTERNAZIONALI

*quaderni di
discussione*

N. 87. 03

Centenario

Dell'International Statistical Institute (ISI):

Adozione delle tecniche informatiche per la statistica

FRANCO MAROZZA

istat

I quaderni di discussione sono a circolazione ristretta e non impegnano la responsabilità dell'ISTAT ma riflettono solo il punto di vista degli autori. Non possono, quindi, essere citati e fatti circolare senza il permesso degli autori.

Le richieste vanno indirizzate a :
«ISTAT - Centro Documentazione - Dr.^{ssa} Borgnino-Valenzano
Via Balbo, 16 - 00100 - ROMA

stat

BIBLIOTECA
DOCUMENTAZIONE

N. 87. 03

Centenario

Dell'International Statistical Institute (ISI):

Adozione delle tecniche informatiche per la statistica

FRANCO MAROZZA

FEBBRAIO 1987

Franco MAROZZA

Vicario del Direttore Centrale

Degli Studi, Contabilità Nazionale

ed Informatica

FRANCO MAROZZA

Centenario
dell'International Statistical Institute (ISI):
adozione delle tecniche informatiche per la statistica

1. Premessa
2. L'Istituto internazionale di statistica
3. La 45^a sessione ISI di Amsterdam svoltasi dal 12 al 22 agosto 1985
4. L'adozione delle tecniche informatiche per la statistica
5. Dettaglio di alcune relazioni presentate alla 45^a sessione relativamente alla materia informatica
 - 5.1. La diffusione dei dati
 - 5.2. Collegamento degli archivi statistici
 - 5.3. L'uso e l'analisi delle banche dati statistici internazionali
 - 5.4. Nuove tecnologie per la raccolta e la memorizzazione delle informazioni
 - 5.5. Nuove apparecchiature e nuove opportunità per l'applicazione di tecniche informatiche alla statistica
 - 5.6. Interfacce ai sistemi statistici per utenti non esperti
 - 5.7. I programmi generalizzati
 - 5.8. Sviluppi recenti dell'uso del PC nei processi statistici con particolare riferimento ai paesi in via di sviluppo
6. Il software statistico presentato ad Amsterdam
7. Conclusioni
8. Bibliografia generale

Franco Marozza (*)

Centenario dell'International Statistical Institute (ISI):
adozione delle tecniche informatiche per la statistica

1. Premessa

Ad Amsterdam, dal 12 al 22 agosto 1985, nella magnifica cornice del Centro internazionale congressi RAI, si è svolta la 45^a sessione dell'Instituto internazionale di statistica (ISI), che è stata anche la sessione del centenario dell'istituzione.

Il convegno ha avuto particolare rilevanza, non solo per la numerosa partecipazione di delegazioni provenienti dai vari paesi e per la importanza delle relazioni e delle memorie presentate, ma anche perchè alcuni temi specifici si riferivano alla utilizzazione della informatica nell'area statistica e nel corso del convegno era stata organizzata una presentazione

(*) L'autore è funzionario in ISTAT dal 1957; ha ricoperto incarichi presso il settore amministrativo ed opera nel campo della informatica da circa trenta anni; ha partecipato come relatore a seminari e congressi nazionali e internazionali; è autore di numerosi lavori di informatica e collabora nel Comitato scientifico di alcune importanti riviste; è membro della Commissione tecnica presso il PGS e di altre commissioni nel settore della P.A.; ha collaborato ad alcuni progetti finalizzati del CNR ed è docente presso la Scuola superiore del Ministero dell'Interno e il FORMEZ. L'autore ringrazia la sig.ra Flora Giordano e la sig.ra M.Teresa Chiarini per la preparazione e revisione dei testi e per le ricerche bibliografiche.

del software statistico per usi professionali e una esposizione del metodo viennese di statistica pittografica (1).

Per quanto riguarda in particolare la presentazione del software statistico che sarà illustrata nel paragrafo 6, erano installati presso la sede RAI alcuni personal computers IBM PC/AT e APPLE Mackintntosh, sui quali potevano essere condotte praticamente le sperimentazioni.

(1) L'esposizione di pittogrammi o ideogrammi era costituita da una selezione di 4000 disegni, preparati tra il 1928 e il 1956 dal grafico GERD ARNTZ, sulla base dell'idea originaria del dr. OTTO MEURATH, sociologo e filosofo viennese, che decise di mettere alla portata di tutti la comprensione di eventi socio-economici, attraverso la statistica pittografica. Tale metodo, un po' trascurato negli ultimi trenta anni, dovrebbe essere rivalutato per la sua praticità e semplicità e per la immediatezza con la quale si possono fornire le comunicazioni in un mondo sempre più basato sulle immagini.

2. L'Istituto internazionale di statistica

L'Istituto internazionale di statistica fu costituito a Londra nel 1885, per facilitare i contatti tra cultori di statistiche nei vari paesi, per migliorare le metodologie, rendere più uniformi e comparabili i dati e richiamare l'attenzione dei governi sulle questioni relative alle rilevazioni statistiche. Ha sede all'AIA e svolge riunioni biennali in città diverse (2).

In applicazione dell'articolo 8a dell'ISI è stata creata l'International Association For Official Statistics (IAOS), che ha lo scopo di promuovere lo sviluppo dei servizi statistici. Di essa sono organi l'Assemblea generale, il Consiglio e il Comitato esecutivo, le Sezioni regionali e da tre anni il GSI Research Centre (che ha tra l'altro lo scopo di sviluppare software specifico in un centro ove sono disponibili data bases) e l'IASC (International Association for Statistical Computing), che ha avuto una assemblea generale il 20.8.85 ad Amsterdam.

(2) F. ZAHN, 50 années de l'Institut international de statistique; J.W.NIXON, A history of the international statistical Institute 1885, 1960.

3. La 45^a sessione ISI di Amsterdam, svoltasi dal 12 al 22 agosto 1985

Come accennato nella premessa, la 45^a sessione dell'Istituto internazionale di statistica si è svolta ad Amsterdam nell'agosto scorso e in quella occasione l'ISI ha anche celebrato il suo primo centenario (3).

Le precedenti sessioni dell'ISI si erano svolte nel seguente ordine, nelle seguenti date e località:

n.°	anno	località	partecipanti
-	1885	Londra	22
1	1887	Roma	92
2	1889	Parigi	66
3	1891	Vienna	81
4	1893	Chicago	79
5	1895	Berna	106
6	1897	St.Petersburg	94
7	1899	Christiania	108
8	1901	Budapest	138
9	1903	Berlino	245
10	1905	Londra	97
11	1907	Copenhagen	100
12	1909	Parigi	149
13	1911	The Hague	125
14	1913	Vienna	166
15	1923	Brussels	119
16	1925	Roma	203
17	1927	Cairo	162
18	1929	Varsavia	169
19	1930	Tokyo	157
20	1931	Madrid	123
21	1933	Mexico	95
22	1934	Londra	113

(3) ISI, The International Statistical Institute (1885-1985).

23	1936	Atene	126
24	1938	Praga	145
25	1947	Washington DC	614
26	1947	Berna	286
27	1951	Nuova Delhi	272
28	1953	Roma	604
29	1955	Rio de Janeiro	267
30	1957	Stoccolma	572
31	1958	Brussels	292
32	1960	Tokyo	318
33	1961	Parigi	513
34	1963	Ottawa	365
35	1965	Belgrado	353
36	1967	Sidney	362
37	1969	Londra	760
38	1971	Washington DC	441
39	1973	Vienna	728
40	1975	Varsavia	710
41	1977	Nuova Delhi	723
42	1979	Manila	587
43	1981	Buenos Aires	722
44	1983	Madrid	879
45	1985	Amsterdam	1.071

Il programma della manifestazione, che ha avuto risonanza internazionale (4), comprendeva lo sviluppo dei seguenti temi:

1. reporting socio-economico: scelta e fonte dei dati, metodi di presentazione;
2. la definizione e la misura della povertà e le disparità temporali e territoriali nella distribuzione del reddito;
3. la statistica in epidemiologia: lo strumento statistico per migliorare la salute;

(4) ISI, Programma finale della Sessione del Centenario, Amsterdam, 1985.

4. aumento dell'utilità della statistica per la pianificazione socio-economica nei paesi in via di sviluppo;
5. le statistiche del lavoro in un mondo in continua trasformazione;
6. alcuni problemi circa la diffusione dei dati, scelta dei media più idonei, delle strategie nella determinazione dei costi e del ruolo degli intermediari;
7. metodologie di interconnessione degli archivi: problema del rispetto della riservatezza (5);
8. l'uso e l'analisi delle basi dati statistici internazionali;
9. la utilizzazione dei modelli per la stima nelle inchieste per campione;
10. nuove tecniche per le statistiche agricole (6);
11. problemi metodologici nelle inchieste longitudinali e nei panels;
12. nuovi sviluppi nel trattamento delle mancate risposte;
13. la diffusione delle informazioni sulla qualità delle inchieste;
14. problemi di organizzazione delle infrastrutture di rilevazione nei paesi in via di sviluppo;
15. nuove tecnologie per la raccolta dei dati (7);

(5) Anche in ISTAT il problema è stato affrontato sulla scia delle esperienze americane, canadesi e inglesi. In uno studio di G. BARCAROLI sono contenute alcune proposte per la realizzazione di procedure per la protezione della riservatezza nell'ambiente ADABAS (banche dati) e in quello delle elaborazioni batch.

(6) F. MAROZZA, Il 3^o Censimento generale dell'agricoltura, in I servizi demografici, n. 12, 1982 e in Comuni d'Italia, 1983.

(7) F. MAROZZA, Il microfilm nel Sistema informativo, in Lo Stato civile italiano, n.3/4, 1979, pagg. 187-192 e 252-256; F. MAROZZA, La macchina che legge, in Lo Stato civile italiano, n. 4, 1979, pagg. 229-231; F. MAROZZA, E. SALERNO, Rilevazione e correzione degli errori dei sistemi di elaborazione elettronica, in Lo Stato civile italiano, gennaio 1981; F. MAROZZA, Il trattamento dei modelli di lettura ottica relativi alla indagine statistica sui trasferimenti di residenza, in Lo Stato civile italiano, n. 7 luglio 1977, pagg. 425-430; F. MAROZZA, E. SALERNO, La raccolta distribuita delle informazioni statistiche, in Comuni d'Italia, n.

16. nuove opportunità e campo di applicazione per le tecniche informatiche in statistica;
17. programmi generalizzati di elaborazione delle rilevazioni;
18. interfacce con i sistemi statistici per i non programmatori;
19. recenti tendenze nell'utilizzazione dei PC nel trattamento dei dati statistici;
20. i modelli semiparametrici;
21. l'analisi fattoriale con dati discreti e continui;
22. teoria multidimensionale dei valori estremi;
23. recenti sviluppi nella classificazione dei dati;
24. statistiche in criminologia;
25. statistiche dell'alimentazione.

Sull'argomento previsto al punto 13 si è svolta anche una tavola rotonda (8), ma su tutti i temi si sono avute interessanti relazioni e vivaci discussioni. Inoltre sono state presentate comunicazioni su invito speciale del Presidente e una selezione dei migliori lavori dei giovani diplomati dei paesi in via di sviluppo (9).

1, 1982; F. MAROZZA, I sistemi di registrazione dati, in Lo Stato civile italiano, n. 11, novembre 1979.

(8) Il comitato d'onore e quello di organizzazione della manifestazione comprendevano personalità politiche olandesi e rappresentanti nel mondo accademico e statistico; erano presenti il presidente dell'ISI J. DURBIN (UK), che ha svolto la relazione introduttiva sul passato e il futuro dell'ISI, i vice presidenti S. MORIGUTI (Giappone), M.A.R. EL-BADRY (Egitto), G. KULLDORFF (Svezia), R.P. MENTZ (Argentina), J.L.NORWOOD (USA) e V. NYTRAI (Ungheria). La delegazione italiana era particolarmente numerosa e composta dai seguenti docenti delle Università indicate tra parentesi: FABRIS e PROVASI (Padova); FERRARI, MORACHIOLI, MARLIANI e PIETRO (Firenze); GIAIMO e MILITO (Palermo); GOLINI, GUARINI e LUCIA (Roma); GOZZI (Parma); RACUGNO (Cagliari); LEMMI (Siena); LOPRESTI (Bari); MARVULLI, MONTINARO e NICOLINI (Torino); MONTANARI (Perugia); MORABITO (Milano); per l'Istat erano presenti il Direttore Generale dr. PINTO, il prof. G.LETI e il dr. A.BUZZI DONATO del CSS, i dottori A.ORSI, F.MAROZZA, M.R.ALVANO, A.CORTESE, B.LEONE, I.SANETTI, A.DE SIMONI.

(9) Tutte le relazioni presentate sono raccolte in 4 volumi e disponibili in ISTAT, unitamente alla documentazione sul software specialistico.

Tutti gli argomenti sono stati di particolare importanza, ma, come informatico, chi scrive ha esaminato con maggiore attenzione i temi di cui ai punti 6 (diffusione dei dati), 7 (collegamento archivi), 8 (le basi dati), 15 (nuove tecniche di data collection), 17 (programmi generalizzati di elaborazione), 18 (interfaccia per gli end users), 19 (i PC in statistica): di tali argomenti si parlerà in dettaglio nei paragrafi 4, 5 e 6.

4. L'adozione delle tecniche informatiche per la statistica.

I cambiamenti che si sono verificati nella tecnologia hanno modificato da tempo l'atteggiamento sia degli utenti che di chi sviluppa i sistemi che elaborano informazioni.

La tecnologia è penetrata in sempre più numerose aree dell'organizzazione e, contemporaneamente, gli utenti hanno preso coscienza dei loro doveri di "possessori" sia dei sistemi che dei dati ed hanno richiesto sistemi migliori, più flessibili e più "amichevoli". Di conseguenza, chi sviluppa i sistemi informativi, è stato costretto a prendere in maggiore considerazione le richieste degli utenti dei sistemi di elaborazione dei dati ed è stata sempre più riconosciuta l'importanza dell'interfaccia tra le macchine e gli utilizzatori.

Per quanto concerne la tecnologia informatica, si deve riconoscere che la statistica è stata una delle prime utilizzatrici dell'informatica: senza voler scomodare un vecchio censimento degli Stati Uniti, quando per la prima volta fu usata la scheda perforata, è noto a tutti che la statistica da sempre ha utilizzato sistemi di elaborazione dati per il trattamento delle informazioni.

Più recente, invece, è l'utilizzo di tali sistemi anche nell'area della gestione delle informazioni, nella conservazione ordinata delle stesse e nella diffusione delle informazioni (10).

Se trenta anni fa la meccanografia di tipo tradizionale trattava prevalentemente dell'acquisizione dati su schede perforate e la macchina più importante del sistema era la selezionatrice, che consentiva di ordinare e contare le frequenze relative alle varie modalità, l'ingresso degli elaboratori e dei più moderni supporti di archiviazione ha modificato il modo di utilizzo dell'informatica ed ha anche determinato

(10) F. MAROZZA, L'automazione negli enti pubblici e il suo impatto con l'utenza, in Informatica e enti locali, n. 3, 1984; F. MAROZZA e altri, Automazione d'ufficio, teledocumentazione, telematica, in Informatica e documentazione, n. 4, 1983.

un riavvicinamento tra statistici e informatici, i cui rapporti non sono stati sempre "ideali".

Il diffondersi dell'informatica distribuita ha anzi consentito all'utente statistico di tornare ad essere indipendente dal tecnico informatico e di riappropriarsi della fase di trattamento dell'informazione, che gli era progressivamente sfuggita di mano. Ma il vero salto differenziale è stato realizzato quando lo strumento ha cominciato a consentire di trattare, oltre ai dati, i testi, le immagini e la voce e si è trasformato anche in strumento di diffusione dell'informazione.

Con tale strumento diventa estremamente più facile e rapida la cattura del microdato di base e, una volta trattato opportunamente tale dato, prevalentemente con pacchetti predisposti e di facile uso, è possibile ottenere tutta una serie di informazioni, le quali - questa è la grossa novità - con lo strumento informatico o parainformatico (videotel, televideo, ecc.) possono essere agevolmente diffuse.

In Italia il progetto di riforma del sistema statistico nazionale è tutto impostato sulla ipotesi di utilizzo delle nuove tecnologie informatiche e telematiche e con la costante preoccupazione di coordinare e standardizzare i diversi sistemi informativi (11).

Il DDL 1108 di riforma, infatti, disciplina le attività di rilevazione, elaborazione, diffusione ed archiviazione dei dati statistici, svolte dalla amministrazione dello Stato e da altre amministrazione ed enti pubblici. La finalità della legge è quella di realizzare l'unità di indirizzo, l'omogeneità organizzativa e la razionalizzazione dei flussi informativi statistici. Vengono anche determinati i criteri per rendere omogenee le attività di rilevazione ed elaborazione statistica compiute nell'ambito del programma e delle iniziative degli altri soggetti del servizio statistico nazionale

(11) Senato della Repubblica, DDL 1108, IX legislatura; F.MAROZZA, Il progetto di riforma del SSN, in Informatica e enti locali, n.3, 1985.

e la omogeneità di rilevazione ed elaborazione non si ottiene se non si fissano degli standards ai quali attenersi rigorosamente.

L'indirizzo più importante della riforma dal punto di vista informatico, è quello che affida all'ISTAT il compito di promuovere e favorire accordi tra i soggetti del Servizio Statistico Nazionale, per l'utilizzazione, ai fini della esecuzione delle rilevazioni, delle reti esistenti. Una tale spinta innovativa non è rintracciabile in altri provvedimenti legislativi, neppure quelli rivolti a regolare il settore specifico delle comunicazioni e trova riscontro solo nel cosiddetto progetto Mirabelli che regola dal punto di vista legislativo la materia informatica (12).

Nell'esercizio della propria potestà organizzativa gli enti ed organismi del Servizio Statistico Nazionale si doteranno, secondo il progetto di riforma dell'ISTAT, anche in forma associata, di strutture tecniche idonee ad assicurare la rilevazione dei dati e la collaborazione con l'ISTAT. I principi affermati sono due ed entrambi di rilevante importanza: con il primo si ribadisce la collaborazione in forme associate dei soggetti del Servizio Statistico Nazionale; con il secondo la necessità di dotarsi di strutture tecniche - e quindi anche informatiche - per il raggiungimento delle finalità statistiche.

Più che gli articoli del DDL 1108, è soprattutto la relazione che è orientata nettamente in senso informatico, in quanto si parla ripetutamente di necessità di sistemi informativi integrati e non ridondanti, di esigenza di sviluppo dell'informatica, che consente l'utilizzo di archivi amministrativi a fini statistici, della creazione di sottosistemi statistici e dei loro collegamenti, di necessità di omogeneizzare i

(12) F. MAROZZA e altri, Le proposte, in L'Informatica nell'amministrazione dello Stato, Atti del 2° Convegno nazionale DC su Situazione e prospettive dell'industria elettronica italiana, 14.6.1978, pagg. 157-163, Ed. 5 lune, Roma; F. MAROZZA, L'automazione negli enti pubblici ed il suo impatto con l'utenza, in Informatica ed enti locali, n. 3, 1984; F. MAROZZA, Il microfilm nel Sistema informativo, in Lo Stato civile italiano, nn. 3/4, 1979.

dati provenienti dalle diverse fonti, di definire un linguaggio di base per la circolazione dell'informazione da utilizzare dai diversi portatori, di armonizzare i vari sistemi informativi amministrativi, della esigenza di favorire il crearsi di sistemi informativi regionali uniformi e, infine, di sviluppare un servizio di diffusione e circolazione delle informazioni più rispondente ai bisogni dell'utenza, basato sui data bases, creando un sistema integrato di informazioni in cui ogni utilizzatore possa facilmente individuare i dati disponibili e necessari alle sue esigenze e possa ottenerli tramite una diversa organizzazione territoriale che avvicini l'offerta alla domanda.

5. Dettaglio di alcune relazioni presentate alla 45^a sessione relativamente alla materia informatica

Come detto nel paragrafo 3, tra i vari argomenti trattati nel corso della conferenza sono stati presi in particolare esame quelli della diffusione dei dati, del collegamento degli archivi, delle banche dati, delle tecniche di raccolta, dell'applicazione di tecniche informatiche per la statistica, dei programmi generalizzati, delle interfacce con gli utenti, dell'uso dei PC in statistica.

5.1. La diffusione dei dati

Tale argomento, che a lungo è stato dibattuto anche in ISTAT (13), è stato trattato da J.W. DUNCAN (U.S.A.), che ha previsto che la tecnologia, in cambiamento costante, avrà senza dubbio una grande importanza per il ruolo degli uffici nazionali di statistica, soprattutto per quanto riguarda la disseminazione e la integrazione delle informazioni. Tale cambiamento pone alcune questioni critiche per gli uffici di statistica e per le organizzazioni statistiche nazionali (14).

Sempre sullo stesso argomento, E. RAPAPORT (15) ha affermato che tutti gli istituti di statistica pongono la loro attenzione per migliorare il servizio agli utilizzatori, presentando una statistica semplice e comprensibile, che viaggia con il progresso tecnologico su nuovi canali di diffusione, tra i quali hanno grande importanza le banche dati, anche se

(13) Lo studio della diffusione ha preso le mosse dallo studio BSP, condotto con la nota metodologia della IBM all'inizio degli anni '80: Analisi e prospettive per la pianificazione dell'ISTAT, 1982; F. MAROZZA, I contratti di utilizzazione in linea degli archivi statistici ufficiali, Atti del 3^o Convegno sull'informatica giuridica, Roma 1983; F. MAROZZA, La diffusione della informazione statistica, in Sistemi e automazione, 1984.

(14) J.W.DUNCAN, Changing responsibilities for dissemination of technology in the modern era.

(15) E. RAPAPORT, The dissemination of official statistics, new approaches, new media, old issues.

la carta stampata continuerà a mantenere la sua importanza. In particolare RAPAPORT parla nella sua relazione dell'importanza dell'informazione sulla informazione (metainformazione), del modo di presentare l'informazione, della qualità dell'informazione e, infine, del tema della determinazione dei prezzi delle pubblicazioni e degli altri supporti: la linea dei bassi prezzi è giustificata, ma è auspicabile un aggiustamento verso l'alto, ipotizzando anche un mercato di prezzi liberi con la diffusione teleinformatica.

5.2. Collegamento degli archivi statistici

Anche questo tema è stato a lungo dibattuto in ISTAT e di fatto sono stati realizzati alcuni successi (16),

(16) Come altri istituti nazionali di statistica, l'ISTAT nello svolgimento della propria attività istituzionale, utilizza dati rilevati direttamente presso il settore delle famiglie e delle imprese e dati che provengono da altri enti e che rappresentano un prodotto secondario di una attività amministrativa. La linea di tendenza a sfruttare le fonti amministrative dovrà essere accentuata notevolmente, secondo i propositi espressi dal Presidente dell'ISTAT al 2° convegno sulla informazione statistica in Italia, tenuto nel 1981 e secondo quanto previsto dalla legge di riforma dell'Istituto, che è prevalentemente incentrata sulla collaborazione statistica tra l'ente centrale e le altre amministrazioni centrali e locali, che debbono avere appositi uffici di statistica e che colloquiano con il centro utilizzando strumenti di tipo informatico. Lo scambio di informazioni tra amministrazioni pubbliche non è certo una novità dei nostri giorni ed in particolare lo scambio di informazioni statistiche è stato facilitato dai progressi dell'informatica che ha reso possibile il trasferimento rapido ed economico di grandi quantità di dati e l'integrazione di archivi prima non coordinati e non comunicanti. Recentemente si è addirittura ipotizzato che i censimenti - indagine statistica per eccellenza - vengano sostituiti con il ricorso ad archivi elettronici. La situazione attuale italiana vede una serie di iniziative ed alcune realizzazioni che saranno passate rapidamente in rassegna e che si riferiscono principalmente alle seguenti aree: contatti con i vari sistemi anagrafici comunali per ottenere i dati relativi al movimento della popolazione direttamente registrati su supporto e per sostituire nel futuro i registri della popolazione ai censimenti decennali; contatti con il sistema tributario per ottenere dati di confronto limitatamente alla struttura anagrafica del mondo delle imprese; contatti con il sistema

mentre in altre aree si deve parlare, se non di insuccessi, almeno di situazioni di stallo (17).

della programmazione sanitaria per ottenere i dati di struttura ed i dati di consuntivo finanziario delle unità sanitarie locali; contatti con il sistema informativo dell'istruzione per ottenere dati di struttura del mondo della scuola; contatti con il sistema giudiziario per coordinare ed integrare i dati sul movimento delle pratiche negli uffici giudiziari con quelli relativi agli organici degli uffici stessi; contatti con il sistema informativo previdenziale per ottenere dati di struttura sull'assistenza e la previdenza; contatti con i vari sistemi comunali per una radicale trasformazione dei metodi di acquisizione dei prezzi; contatti con il sistema delle anagrafi delle imprese (CERVED) per determinare con migliore approssimazione lo stato aggiornato delle imprese in Italia; contatti con i vari sistemi informativi regionali, da un lato nel tentativo di assumere direttamente le informazioni di base a livello territoriale e dall'altro per tentare di realizzare un flusso informativo di ritorno regionale; infine esistono una serie di altri progetti e studi che si riferiscono all'assunzione diretta di informazioni relative al sistema del lavoro, al sistema delle dogane (commercio con l'estero), al sistema dei trasporti, al sistema delle opere pubbliche ed altri.

(17) Naturalmente i rapporti tra le varie amministrazioni non sono sempre facili e semplici, in quanto spesso intervengono particolari problemi relativi alla competenza specifica e spesso esclusiva delle amministrazioni stesse. Fino a qualche anno fa sono stati di una certa rilevanza i problemi della compatibilità hardware che hanno limitato la circolazione dei dati e tutt'oggi sono di ostacolo alcuni problemi software che impediscono la portabilità dei programmi. Esistono ancora problemi di riservatezza e assumono particolare importanza, in funzione del fatto che i vari archivi integrati tra loro possono dare alla informazione un valore aggiunto tale che possa essere violata la sfera di riservatezza dell'individuo. Esistono, infine, problemi di armonizzazione dei concetti e delle definizioni, senza i quali gli archivi non possono essere integrati tra loro. Per risolvere i problemi che nascono nella ripartizione del lavoro statistico tra le varie amministrazioni, sono state individuate tre strade. La prima, più generale, è quella legislativa: come detto, un disegno di legge è stato presentato dal Governo al Parlamento per organizzare un sistema di statistiche ufficiali coordinato da un Consiglio nazionale di statistica. La seconda, più specifica, è quella delle convenzioni, una sorta di contratti tra ISTAT e singole amministrazioni per costruire sistemi informativi nei domini di comune interesse: ne è un esempio un accordo con il Ministero del Lavoro. La terza strada è quella del coordinamento tecnico guidato da commissioni di studio: un esempio è fornito dalle statistiche del settore statale, che sono state riorganizzate da una Commissione istituita dal Ministero del Tesoro

Ad Amsterdam il tema è stato affrontato da quattro relazioni (18), che, però, hanno prevalentemente trattato dei problemi del rispetto della riservatezza (19), e del diritto alla protezione della vita privata, che sono associati al problema del collegamento delle informazioni raccolte da un ufficio governativo per ragioni statistiche o di ricerca. La relazione di L. ALEXANDER tratta del collegamento di archivi di microdati contenenti informazioni individuali provenienti da fonti diverse e collegabili attraverso elementi comuni di identificazione: tale collegamento solleva seri dubbi sulla legittimità di tale operazione e impone l'utilizzo di necessarie cautele, per impedire gli abusi ed assicurare il rispetto della riservatezza, soprattutto quando gli archivi provengono da amministrazioni differenti.

L.H. COX e R.F. BORUCH (USA), hanno messo in evidenza l'utilità e il valore delle tecnologie di linkage per migliorare la qualità, la completezza e l'utilità dei dati, spesso a costi bassi in termini di risorse organizzative e impegno dei rispondenti. G. PAASS (20) ha illustrato, con considerazioni teoriche e simulazioni empiriche, lo stato dell'arte e le future prospettive dello statistical matching methods (S.M.). Le perturbazioni di accoppiamento di dati con

con la partecipazione di esperti dell'ISTAT e della Banca centrale. I risultati finora conseguiti non sono certamente confortanti sotto tutti i profili: se ci si dovesse fermare alle prime difficoltà, certamente alcuni dei più importanti progetti sarebbero stati già abbandonati, ma all'Istituto Centrale di Statistica si è fermamente sicuri che il ruolo che la legge e le tradizioni assegnano all'ISTAT debbono essere assolti ripartendo le responsabilità tra i vari organi della rete informativa nazionale coordinati dall'organo centrale.

(18) L. ALEXANDER, La liaison des données statistiques et son effet sur le droit à la protection de la vie privée; P. JENSEN, L. TRYGESEN, Linkage of records on objects of different kinds and methodological problems and practical experience; L.H. COX e R.F. BORUCH, Emerging policy issues in record linkage and privacy.

(19) CALOSSO, La libertà di informazione e la riservatezza dei dati personali, relazione ISTAT presentata alla 31^a riunione dei D.G. degli uffici di statistica.

(20) G. PAASS, Statistical record linkage methodology: State of the art and future prospective.

differenti valori di x possono determinare una distorsione della distribuzione xy nel caso del mancato rispetto della ipotesi della indipendenza condizionale (CIA), che impedisce una riproduzione corretta della distribuzione yz .

Il documento di P. JENSEN e L. TRYGESEN, danesi, considera, infine, alcune esperienze di linkage di records provenienti da differenti fonti: per esempio informazioni relative ad una persona (età, residenza, indirizzo), con informazioni relative al suo posto di lavoro (indirizzo, attività). Il documento tratta anche dei problemi che si sono affrontati in Danimarca, soprattutto per quanto concerne la creazione della chiave di collegamento (21).

5.3. L'uso e l'analisi delle banche dati statistici internazionali.

Su questo argomento chi scrive ritiene importanti i contributi presentati alla conferenza di Amsterdam da R.O. BARTRAM e S. BAUM e T.D. KOUASSI (22).

Nel primo lavoro i due autori affermano che, mentre l'attività degli istituti nazionali è rivolta all'interno del paese, quella delle agenzie internazionali è rivolta all'estero. Fa eccezione l'U.S. Bureau of Census, che raccoglie e diffonde statistiche di altre 200 nazioni, statistiche ovviamente residenti su data bases.

Il documento di KOUASSI descrive, invece, la banca dati statistici della CEA (uno dei componenti del sistema panafricano di documentazione e d'informazione), che usa una delle tecnologie più recenti per trasmissione dei dati contenuti negli archivi, che sono organizzati su tre livelli. Il primo fornisce pochi dati di base su tutti i paesi africani

(21) F. MAROZZA, I registri della popolazione e la loro utilizzazione a fini statistici, in *Lo Stato civile italiano* n.5, 1979, pagg. 295-302.

(22) R.O. BARTRAM e S. BAUM, An international data base at a national statistical agency; T.D. KOUASSI, The ECA statistical data bank.

(2/3 pagine); il secondo contiene la maggior parte delle informazioni dettagliate relative ai vari paesi; il livello tre contiene le serie cronologiche dettagliate.

5.4. Nuove tecnologie per la raccolta e la memorizzazione delle informazioni

Sono stati selezionati tre documenti che chi scrive ritiene molto interessanti: il primo è dell' U.S. Census Bureau (23) e tratta del CATI (Computer Assisted Telephone Interviewing), che impiega interattivamente i sistemi di elaborazione per la raccolta telefonica delle interviste.

L'introduzione massiccia di tale sistema è in uso nei settori commerciali, accademici e governativi e le esperienze quotidiane sono discusse sotto il profilo dei costi HW e SW, quello della velocità di produzione e, infine, sotto il profilo della qualità dei dati.

Il secondo lavoro è di L. LYBERG (24) e descrive la medesima esperienza in Svezia, ma si sofferma non solo sulle attuali realizzazioni, ma sui piani futuri rivolti, non alla raccolta centralizzata, ma agli intervistatori sul campo muniti di PC portatili, utili per le interviste telefoniche, domiciliari e per il data-entry.

Infine, il lavoro di M.E.J. BEMELMANS-SPORK e D. SIKKEL parla delle esperienze olandesi nella raccolta dati con l'utilizzo di PC portatili. Tre sono i principali quesiti cui gli esperimenti dovrebbero rispondere: sono gli intervistatori capaci di fare il loro lavoro con tali strumenti? i rispondenti sono disposti a rispondere se si utilizzano tali mezzi? è possibile scrivere un SW idoneo a tale funzione? Le risposte sono in genere positive (25).

(23) The status of computer assisted telephone interviewing: 1985, by W.L.NICHOLLS II, U.S. Census Bureau e R.M. GROVES, University of Michigan.

(24) L. LYBERG, Plans for computer- assisted data collection at Statistics Sweden.

(25) M.E.J. BEMELMANS-SPORK, D. SIKKEL, Data collection

5.5. Nuove apparecchiature e nuove opportunità per l'applicazione di tecniche informatiche alla statistica

Su questo argomento sono stati predisposti numerosi documenti, la prevalenza dei quali ha affrontato il problema posto dalla introduzione delle nuove tecniche informatiche, che impongono una trasformazione del rapporto tra informatici e statistici e la determinazione di nuove regole del gioco (26).

Il documento di F. KRAMARZ, Y. THIRY e F. VENNAT (27) dell'INSEE, prende lo spunto per trattare tale problema dallo studio di una nuova struttura dei dati contabili in occasione della determinazione della nuova base dei conti nazionali (28).

with hand-held computers.

(26) F.MAROZZA, Nuovi orientamenti sull'utilizzazione degli elaboratori elettronici nella informazione statistica, in Atti del Convegno sull'informazione statistica in Italia, Annali di Statistica, Serie VIII, Vol.26, 1971, pagg.239-252; F.MAROZZA, I servizi statistici nei prossimi dieci anni, in Informazioni, luglio-agosto 1980, bollettino di documentazione FS; C.VITERB C, F.MAROZZA, Technical and Organizational Innovations in the Informatic Statistical Field in Coming Years, in Statistical Services in Ten Years' Time, ONU, Pergamon, 1978; F.MAROZZA, I servizi statistici nei prossimi dieci anni, in Informatica e documentazione, n.1/80; F.MAROZZA, Il Sistema informativo statistico ed i nuovi strumenti di elaborazione elettronica, ISTAT 1979, Collana d'informazioni n.4, pagg. 4-44; F.MAROZZA, Le innovazioni tecniche prevedibili nel campo della statistica nel prossimo decennio, in Informatica e documentazione, rivista dell'Inforav, Istituto per lo sviluppo e la gestione avanzata dell'informazione, n. 2, giugno 1977 pagg. 103-111; F.MAROZZA, Gli elaboratori elettronici nell'ISTAT, relazione per il 50° dell'Istituto Centrale di Statistica, in Cinquanta anni di attività, 1926-1976: cap. 5, Gli elaboratori e l'informazione statistica automatizzata (pagg. 46-62); Appendice C, Configurazioni degli elaboratori ISTAT (pagg. 383-396); Appendice D, I linguaggi e i sistemi operativi (pagg. 397-402).

(27) F. KRAMARZ, Y. THIRY e F. VENNAT, Le système informatique de la nouvelle base des comptes nationaux: la définition de nouvelles règles de représentation des grandeurs économiques par un travail commun des comptables et des informaticiens.

(28) In ISTAT, la situazione del SW deve essere razionalizzata e organizzata, in quanto ad un periodo di scarsa attenzione per il SW esistente è subentrata una fase di acqui-

Il documento di BO SUNDGREN (29) fa la storia dell'impatto avvenuto negli ultimi 20 anni nell'Istituto di statistica svedese nell'utilizzazione del SW generalizzato. I nuovi metodi di strutturazione dei dati e di progettazione dei sistemi sono i requisiti di una evoluzione che si è tradotta in un aumento di produttività ed efficienza. Questa evoluzione ha modificato la funzione degli informatici, favorendo la decentralizzazione del DP e restituendo agli statistici il controllo integrale della produzione statistica.

Questo anche perchè nell'Istituto svedese la situazione del SW era divenuta complessa e troppi differenti prodotti erano usati in parallelo, con l'aumento dei costi di acquisizione, aggiornamento, sviluppo e manutenzione: sono stati quindi raccomandati alcuni prodotti che consentono delle efficaci interfacce utenti. Negli anni futuri si progetta di far ricorso ai sistemi di intelligenza artificiale - sistemi esperti della quinta generazione - per mettere a punto dei sistemi di produzione statistica ancora più efficienti.

Il documento di D. DI-QING (30), dà un profilo inedito della situazione del DP statistico nella Repubblica popolare cinese a partire dal 1960. L'ente produttore di statistiche ufficiali dal 1974 ha il proprio elaboratore fatto in casa, anche se vengono usati attualmente elaboratori IBM 43XX; vengono usati pacchetti standard internazionali e la decentralizzazione del DP è una delle principali caratteristiche della riuscita del censimento 1982. I 29 centri elettronici del paese sono riuniti nell'ISIS, ma il paese deve ancora molto imparare dalle esperienze straniere. Nel documento di HIRCHIKO

sizione e utilizzo disordinato di parecchi SW di base e di differenti SW statistici; F. MAROZZA, Statistica ufficiale e informatica, in Stato e regione, n. 6, novembre-dicembre 1980; F. MAROZZA, La formazione informatica degli statistici, in I servizi demografici nn. 7/8, 1982; F. MAROZZA, Informatica e statistica, relazione al 5° convegno di PADOVA, 1/4 Aprile 1985.

(29) B. SUNDGREN, How to satisfy a statistical agency's need for general survey processing programs.

(30) D. DI-QING, Data processing for statistical surveys in China.

KOYAMA (31) è riportata, invece, l'esperienza giapponese nell'area del DP statistico, che risale anch'essa al 1961 e, attraverso le varie generazioni di elaboratori IBM, arriva ai 3083 del 1984, alternando però l'uso di apparecchiature NEAC fatte in casa. Il documento descrive, infine, il sistema in linea per la correzione dei dati e il sistema automatico di esame dei risultati delle indagini.

5.6. Interfacce ai sistemi statistici per utenti non esperti

Il documento inglese di C.DAFFIN, P. DUNCOMBE e di G.B. WETHERILL (32), descrive l'esperienza degli autori nel corso della preparazione di un pacchetto d'analisi statistica per utenti non specialisti e l'altro documento di D.J. HAND (33) tratta della scelta di una tecnica statistica per guidare il ricercatore; due grandi classi di programmi sono prese in esame: quelli basati su strutture esplicite di alberi di decisioni e quelli basati su una architettura di sistemi di produzione. Il documento di J.L. ROOS (34) dell'INSEE, considera i DB statistici, nei quali è difficile reperire e trattare l'informazione che è molto numerosa: per affiancare i SW statistici tradizionali (TSP, SAS, BMDP), viene proposto un sistema informativo implementato in LIPS, che permette di gestire contemporaneamente tavole di dati e nomenclature, attraverso una rappresentazione orientata all'oggetto.

Il documento di A.R. WILKS (35), propone alcune idee di menu gerarchici e/o a finestre, che possano effettivamente aiutare l'interattività tra uomo e macchina, mentre il lavoro

(31) H. KOYAMA, Efficient utilization of computers for reliable statistical data processing.

(32) C. DAFFIN, P. DUNCOMBE e G.B. WETHERILL, A user friendly survey analysis program.

(33) D.J. HAND, Choice of statistical technique.

(34) J.L. ROOS, Recherche et traitement des données dans une base de données statistique orientée objet.

(35) A.R. WILKS, Simple and rapid control through dual-loop hierarchical menus.

di J. DÖRNYEI (36), propone l'ideazione di un progetto pilota per un sistema esperto, con l'integrazione di una base di dati conversazionali, un sistema di metainformazioni e un pacchetto di analisi di dati.

Infine, il lavoro di A. PRAT, M. MARTI e J.M. CATOT (37), presenta lo STATXPS, un'interfaccia tra l'utente non esperto e il sistema statistico, con le caratteristiche di un sistema esperto di uso particolare per le serie temporali.

5.7. I programmi generalizzati

L'utilizzo di programmi generalizzati nell'area statistica non è certo una novità e praticamente l'ISTAT ha avuto la possibilità negli ultimi anni di esaminare e utilizzare i più importanti programmi generalizzati disponibili e predisposti dai vari istituti di statistica.

In particolare sono utilizzati in ISTAT:

- TPL : programma generalizzato di preparazione di tabelle statistiche, prodotto e diffuso dal Bureau of Census USA.
- AERO : sistema di controllo e correzione automatica errori, prodotto dall'Istituto di Statistica ungherese.
- CONCORD : prodotto simile al precedente, che è stato studiato in ISTAT in maniera di integrarlo con AERO, producendo AEROCON, che riunisce i vantaggi dei due singoli prodotti.
- SAS : pacchetto per le elaborazioni statistiche, commercializzato in Italia dalla Syntax e comprendente anche un modulo grafico.

(36) J. DÖRNYEI, Central Statistical office of Hungary, From dialogue systems to expert systems: conception for application of AI in statistical information management.

(37) A. PRAT, M. MARTI, J.M. CATOT, An interface between a non expert user and statistical systems: the statxps.

- SPEA EASY: pacchetto per la gestione delle serie e dei modelli econometrici, prodotto dal MIT.
- RAPED : sistema gestionale dell'Istituto di Statistica canadese, basato su un DB di tipo relazionale in ambiente MVS.

5.8. Sviluppi recenti dell'uso del PC nei processi statistici con particolare riferimento ai paesi in via di sviluppo

Il documento di T. WALCZAK (38), del Central Statistical office of Poland, descrive le linee generali della introduzione dei PC in statistica e nello stesso tempo sottolinea i problemi indotti (39).

V.K.SAMARANAYAKE e S. ABEYASEKERA del Dipartimento di Statistica e Informatica della Università di Colombo (SRI LANKA), hanno sottolineato nel loro documento di Amsterdam (40), il successo della diffusione del PC per l'analisi dei dati, soprattutto nella ricerca in agricoltura. Il documento parla anche dell'assistenza ricevuta dall'Università di Reading (UK) nello sviluppo del pacchetto di software statistico INSTAT.

Nel documento di B.N.DISKIN, dell' U.S. Bureau of Census (41), si parla della introduzione dei microcomputers negli uffici statistici dei paesi in via di sviluppo; a differenza di quanto avviene per il grande computer, dove le risorse sono limitate, i modelli installati non sono i più recenti, le apparecchiature spesso non compatibili e il SW piuttosto specialistico e di tipo professionale, i minicomputers hanno

(38) T. WALCZAK, Micro-computers in statistics.

(39) F. MAROZZA, Informatica distribuita per la Statistica, Seminario ISTAT, Ministero Industria Commercio Artigianato, UNIONCAMERE, Roma 23/24.5.1984; F. MAROZZA, Informatica distribuita per la statistica, in Informatica ed enti locali, n. 4 dic. 84.

(40) V.K. SAMARANAYAKE, S. ABEYASEKERA, Development of statistical computing on microcomputers at the Colombo University.

(41) B.N. DISKIN, Microcomputers in developing country statistical offices: current use and a look to the future.

avuto un utilizzo più appropriato e se ne prevede un ampio sviluppo, soprattutto per le prestazioni a basso costo e per la diffusione di pacchetti di facile utilizzo.

Infine, nel documento di W.J. DIXON e P. BRITT, dell'UCLA USA (42), si parla della rivoluzione che il microcomputer ha portato nell'area statistica, più importante di quella determinata dalla originaria introduzione dei grandi calcolatori. Quattro sono i fattori di tale sviluppo:

- il basso costo dell'HW
- la facilità di uso del SW
- la divulgazione della cultura informatica
- lo sviluppo di SW molto sofisticato.

(42) W.J. DIXON e P. BRITT, *The future of statistical analysis: small computers and the developing countries.*

6. Il software statistico presentato ad Amsterdam

Come detto precedentemente, nel corso del Seminario di Amsterdam, si sono svolte dimostrazioni su microcomputers di elaborazione dati con pacchetti statistici generalizzati (43).

Tra i software presentati è stato possibile visionare i seguenti:

- SPSSTABLES: tabular reporting SW, che produce tavole statistiche di buona presentazione e qualità (44);
- SPSS/PC: SW statistico per PC, da utilizzare velocemente e in maniera interattiva con funzioni grafiche;
- SPSS (Information analysis system): è un SW per mainframes e supermini, comprensivo di un pacchetto di procedure statistiche (analisi multivariate, modelli lineari, varianza, modelli di regressione e rappresentazioni grafiche); tale SW è da anni utilizzato in oltre 3.000 installazioni in 80 paesi su 40 differenti tipi di macchine, da centinaia di migliaia di persone;
- BMDP (Statistical SW), che ha decine di programmi per la costruzione di istogrammi, tavole di frequenza, analisi varianza, covarianza, regressione, cluster analysis, ecc.: è utilizzato nell'area commerciale, industriale, universitaria e governativa (45); i pacchetti possono interfacciare il SAS, SIR e P-STAT e sono utilizzabili su AMDAHL, IBM (mainframe e PC), ICL (Serie 2900), SPERRY (Univac 1100) e HONEYWELL;
- NAG (Numerical Algorithms Group limited): ha una parte di

(43) Vedi S.S.N. (Statistical Software Newsletter), che è il periodico del GESELL SCHAFT FUER STRAHLEN-UND UMWELTFORSCHUNG MBH e dell'IASC.

(44) SPSS Benelux B.V.P.O. Box 111 4200 AC GORINCHEM (Netherlands).

(45) STATISTICAL SW LTD, Cork Farm Centre Demehy's Cross Cork Ireland.

pacchetti dedicati alle funzioni basilari della statistica (46); tra questi erano disponibili nella dimostrazione di Amsterdam:

- a) TSA (an interactive package for time series analysis and signal processing);
- b) MLP (maximum likelihood program);
- c) GLIM, sviluppato per la Royal Statistical Computing Society (generalized linear models);
- d) GEMSTAT (a general statistical program).

Indipendentemente dal materiale presentato ad Amsterdam, si è tentato di completare la ricerca in un mercato in pieno sviluppo e così espanso come quello del software statistico nel quale è difficile condurre un'indagine esaustiva e tanto meno per campione, poichè non è noto l'universo dei produttori (47).

La ricerca citata nella nota precedente ha preso in esame i packages, citati in articoli specializzati e nelle pubblicità apparse nelle maggiori riviste statistiche internazionali, pervenendo così ad una lista di 120 packages o programmi che certamente rappresentano un vasto segmento del mercato se non quasi la sua totalità.

In genere, nella descrizione del SW, si è cercato di individuare l'HW richiesto, le interfacce utenti e le funzioni statistiche previste (tabulazioni, rappresentazioni grafiche, distribuzioni di probabilità, test, varianza, analisi non parametriche, correlazione, regressione, analisi multivariate, serie temporali, ecc. i).

Per quanto concerne il tipo di macchina impiegato, si può affermare che l'offerta di software è ancora largamente condizionata dalla presenza sul mercato di macchine basate su microprocessori a 8 bit: tra queste il primo posto spetta ai

(46) NAG Central office, Mayfield House, 256 Bassbury Road, Oxford OX27 BR.

(47) La maggior parte delle informazioni che seguono sono tratte da uno studio condotto da P. COSTA della Sperry.

modelli Apple dichiarati in oltre il 40% dei packages, mentre il software implementato sull'IBM PC e compatibili sotto MS/DOS rappresenta meno del 40% del mercato. Il software che "gira" sull'IBM PC e compatibili passa al 50% circa se si considera anche quella parte del software che è stato sviluppato in ambiente CP/M 86. Il linguaggio più diffuso è il BASIC, che ricorre in circa il 70% dei packages, seguito dal FORTRAN (10%), dal PASCAL e dall'ASSEMBLER.

I prezzi sono molto variabili e dipendono in larga misura dalle caratteristiche differenziali del SW. Per il 35% dei packages o programmi, il costo è inferiore ai 100\$; per il 25% è compreso tra 100 e 250\$; per il 30% tra 250 e 500\$ e per il 10% va oltre 500\$.

Tra gli altri SW statistici esaminati hanno rilevanza i seguenti:

- IMSL è una libreria di programmi basati su sottoprogrammi FORTRAN, per risolvere problemi di matematica (equazioni differenziali, interpolazione, equazioni lineari e non lineari, matrici) e statistici (regressione, varianza, serie storiche, generazione casuale di numeri, ecc.) (48);
- MINITAB sistema statistico interattivo, adatto per statistici e non statistici, dotato di comandi con facile sintassi e di un elementare modulo grafico (49);
- TSP (time series processor), per la manipolazione di modelli econometrici (50);
- STATPAK (Statistical analysis package) questo pacchetto, che è distribuito in Italia dalla CORTIS/LENTINI di Torino, è stato presentato nel numero di gennaio di ZERO UNO: si tratta di una libreria di cinquanta funzioni statistiche (calcolo probabilità, variabili

(48) IMSL Sales Division NBC Building, 7500 Bellaire Boulevard, HOUSTON 77036 USA.

(49) MINITAB Project, Pennsylvania University 215 Pond Laboratory, University Park Pennsylvania 16802 USA.

(50) TSP International, 204 Junipero Serro Boulevard Stanford, California 94305.

- statistiche, distribuzione funzioni, regressione e correlazione, analisi della varianza) (51);
- APLM pacchetto interattivo per funzioni statistiche scritto in APL (52);
 - SURVTAB (Survival tables), scritto in BASIC (53);
 - LOLITA (Log linear modelling in tables), scritto in Pascal (54);
 - XECUVISION, utilizzato soprattutto per la creazione di immagini e rappresentazioni grafiche da riportare anche su slides (55).
 - PCS (Personal computer statistics), che è un pacchetto preparato in Italia per risolvere problemi di statistica (56).

La ricerca di P. Costa, precedentemente citata, conclude con una tabella di selezione del SW statistico per un PC Sperry:

(51) Walonick Associates, 5624 Girard Avenue S. Minneapolis MN 55419 USA.

(52) N.C.LAURO, Dipartimento matematica e statistica, Università di Napoli, Via Partenope 36, 80121 NAPOLI.

(53) Statistical Computing Consultants.

(54) B.P. Murphy Raine, Medical Statistics Unit, University of Western Australia Nedlands, 6009.

(55) Visual Communications Metusrk, 23 Main Street, Suite 324 Cambridge Ma 2142 USA.

(56) Il pacchetto, prodotto dalla TT SOFT di Milano, è stato presentato nel numero 49 di ZERO UNO (febbraio 1986) e consente il calcolo di indici statistici, l'analisi campionaria di regressione semplice e multipla, della varianza, delle tabelle di contingenza, distribuzioni di probabilità continue e discrete, numeri indice di prezzi, quantità e valore, ecc.

Class	Name	Scope
GENERAL STATISTICAL SYSTEMS	STATPAC	Elementary data analysis
	MINITAB	Teaching Statistics
	BMDP	Experimental data analysis
	SPSS	Survey data analysis
	IMSL	Statistical & Mathematical subroutines' system

SPECIAL URPOSE PAC AGES and PROGRAMS	TSP	Time series and econome- trics
	LOLITA	Multivariate contingency tables analysis
	GLIM	Linear models
	APLM	Multivariate analysis
	SURVTAB	Survival analysis

Per dare una misura dello sviluppo dei packages per l'area statistica si riportano i risultati di una ricerca - per sfortuna non troppo recente - descritta in un rapporto della SISDO, BUTLER COX FUNDATION (57), che per il Regno Unito ha esaminato i packages disponibili, classificati secondo il tipo di applicazione e il tipo di computer:

(57) The BUTLER FUNDATION, Packages applicativi, Rapporto 22, 1981.

Tipo di applicazione	Tipo di computer				
	Mainframe		Minicomputer		Microcomputer
	Prima del 1974	1975-79	Prima del 1974	1975-79	1975-79
Contabilità finanziaria	49	46	58	228	50
Amministrazione	42	48	15	70	12
Vendite e marketing	11	11	4	55	12
Gestione del personale	13	29	25	59	15
Gestione di progetto	17	24	9	27	7
Pianificazione e controllo della produzione	33	46	24	97	19
Distribuzione	1	10	5	23	4
Statistiche	14	23	9	9	10

7. Conclusioni

Nei paragrafi precedenti è stata illustrata, in rapida sintesi, la parte del seminario di Amsterdam che ha interessato in particolare l'informatico.

Nella sessione celebrativa del centenario dell'ISI si è simbolicamente rievocato anche il centenario della introduzione su larga scala dell'uso della scheda meccanografica e - indirettamente - il centenario della nascita delle tecniche informatiche, che tanta parte hanno avuto nella sostanziale modifica delle tecniche statistiche, non solo per quanto riguarda la elaborazione, ma anche per quanto riguarda le fasi di acquisizione e diffusione dei dati. Con l'aver passato in rassegna alcune delle relazioni presentate al convegno, non si pretende certamente di aver dato una informazione completa di quanto ad Amsterdam si è discusso in due intense settimane di lavoro.

Quella fornita è soltanto una selezione bibliografica, che potrà consentire, ai ricercatori interessati, di accedere direttamente ai documenti di base provenienti da esperti di tutto il mondo. Lo stesso discorso vale per la presentazione del software statistico disponibile: si tratta solo di elenchi sommari e non completi, che possono fornire indicazioni per ricerche più approfondite su documenti in parte disponibili presso le biblioteche specializzate e in parte accessibili direttamente presso i produttori.

Nel 1987 la 46^a sessione dell'ISI si terrà a Tokyo.

Dall'agenda provvisoria distribuita, i temi per la 46^a sessione dell'Istituto Internazionale di Statistica a Tokio 1987, sono i seguenti:

A) Topics

1. The trend of official statistics. Some problems connected with the future
2. Development of statistics in Japan. Past experiences and present problems

3. Analysis on historical records. Comparative historical demography
4. Measurement and utilisation of health related indicators
5. Statistical techniques for improvement of quality and productivity
6. Papers by special invitation of the president
7. Best papers by recent graduates from developing countries
8. Measuring change from survey
9. Master sample in national household surveys
10. Variance estimation with complex sample
11. The future of the population census
12. Methodological issues in survey for price indexes
13. Effects of mode of data collection on survey response
14. Statistical education in government and industry
15. Artificial intelligence and expert systems in statistics
16. Statistical graphics
17. Problems of and approaches to analysing large data sets
18. Statistical needs of developing countries
19. Recent developments of software systems for time series analysis
20. Advances in data banks and network. Building and management
21. Statistical methods for spatial image analysis
22. Recursive procedures for time sequenced data
23. New developments in time series modelling and analysis
24. Estimation from accelerated tests and short term tests in reliability and medicine
25. Self similar processes
26. Principles and strategies of data analysis
27. Survey processing on microcomputers
28. Statistical computing in the 90's. Impact of the fifth generation
29. Small area statistics
30. New developments in the statistical methodology in chemistry

B) Special contributed paper topics

1. Current issues in the principles of statistical inference
2. Statistical methodology for studying the status of women
3. Census evaluation
4. Handling missing data in survey
5. Stochastic differential equations. Theory and applications
6. Repeated measurement and crossover design
7. Statistical methods for data with possible outliers
8. Timeliness and accuracy in preliminary estimates and revision of economic data.

Mi si consenta una considerazione conclusiva: chi scrive ha molto apprezzato il valore scientifico e culturale del materiale presentato nella sessione di Amsterdam e, soprattutto, ha avuto modo di constatare la preparazione degli statistici e degli informatici presenti, ma si deve alla fine concludere che i problemi che affliggono il mondo statistico italiano sono gli stessi degli altri paesi e le soluzioni - anche tecnologiche - che vengono adottate in Italia non sono certamente inferiori, nè meno brillanti, di quelle che vengono adottate nel resto del mondo.

Franco Marozza

ISTAT, febbraio 1986

8. Bibliografia generale

W.ALVEY, F.AZIZ, Mortality Reporting in SSA Linked Data: Preliminary Results, Social Security Bulletin, November, 1979.

K.ARNOLD, Screen Updating and Cursor Movement Optimization: A Library Package, Dept. of Computer Science, University of California, Berkeley, 1981.

R.S.BARR, J.S.TURNER, Microdata File Merging through Large-Scale Network Technology, Mathematical Programming Study, Vol. 15, 1-22, 1981.

R.A.BECER, J.M.CHAMBERS, An Interactive Environment for Data Analysis and Graphics, Wadsworth 1984.

E.BENDER, Small Change, in Computerworld, December 31, 1984/January 7, 1985.

E.BJÖRK, The Concept of Base Operators: An Extension to the Relational Algebra towards Statistical Applications, Workshop on the Statistical Computing Project, Budapest, Hungary, pp.12-17 November 1984.

S.BLAKEY, Around the World in 16k, in Computerworld, December 31, 1984/January 7, 1985.

J.BROTH, W.MUDRY, Quality control. Its role in quality assurance and its relation to survey operations, Internal report, Statistics Canada, 1981.

Creative Computing Benchmark, in Creative Computing, July, 1984.

P.CAREY, The Future of the Micro: Looking Ahead at the Next Decade, in Popular Computing, January, 1985.

T.A.CARGILL, Pads Programming Guide, AT&T Bell Laboratories Technical Memorandum # 11276-841015-07TMS, 1984.

J.MC CARTHY, Recursive Functions of Symbolic Expressions and Their Computation by Machine, Communication ACM, April, 1960.

J.M.CHAMBERS, Some Thoughts on Expert Software, Proc. Comp Science and Statistics, 13th Symposium on the Interface, 1981.

J.M.CHAMBERS, D.PREGIBON, E.R.ZAYAS, Expert Software for Data Analysis: An initial Experiment, Proceedings of the 43rd Session of ISI, Buenos Aires, 1981.

C.Z.F.CLARK, J.L.COFFEY, How many people Can Keep a Secret? Data Interchange Within a Decentralized System, American Statistical Association, Proceedings, Washington, D.C., 1983.

L.DANIELSSON, P.A.MÄRSTAD, Prov med handhållen terminaldator/fältdatainsamling i KPI, DATI-projektet, Memo, Statistics Sweden, 1982.

R.DAVIS, Consultation, Knowledge acquisition and instruction: a case study, in Artificial intelligence in medicine, P. Solovits, pp. 57-58, Colorado, Westview Press, 1982.

R.DAWSON, Effective dissemination and marketing of statistical information: productive partnerships with data intermediaries.

F.T.DE DOMBAL, D.J.LEAPER, J.R. STANILAND, A.P.MACCANN, J.C. HORROCKS, Computer-aided diagnosis of acute abdominal pain, Brit.Med.J.2, pp.9-13, 1972.

B.DIXON, B.STUART, L.ISLOW, Considerations for Use of Microcomputers in Developing Country Statistical Offices, 1984.

E.J.DUDEWICZ, V.S.TANEJA, New directions in multivariate statistical analysis and statistical computing.

R.DUFFY, Lotus Is Still on Top After Software War, in PC Week, December 25, 1984.

IP.FELLEGI, D.HOLT, A systematic approach to automatic editing and imputation, Bull Int Statist Inst, 45(1), pp.364-376, 1973.

IP.FELLEGI, Automatic editing and imputation of quantitative data, Proceedings of the 40th Session (4 vols), Bull Int Statist Inst, 46(3), pp.249-253, 1975.

IP.FELLEGI, D.HOLT, A systematic approach to automatic edit and imputation, J Amer Statist Ass, 71(353), pp.17-35, 1976.

D.FRAZIER, Open Architecture Question to Dominate Industry in '95, in PC Week, December 25, 1984.

J.FRIEDMAN, M.A.FISHER ELLER, J.W.TUKEY, An Interactive Multi-dimensional Data Display and Analysis System, Proc. Fourth International Congress for Stereology, PRIM-9, 1974.

G.FURNAS, The FISHEYE view: a new look at structured files, AT&T Bell Laboratories Technical Memorandum # 82-11221-22, 1982.

H.GALLAIRE, J.MINIER, J.M.NICOLAS, Logic and Databases, a deductive approach, ACM Computing surveys, vol 16, num 2, June, 1984.

T.D.GARVEY, J.D.LOWRANCE, M.A.FISCHLER, An inference technique for integrating Knowledge from disparate sources, Proceedings of 7th International Joint Conference on Artificial Intelligence, Vancouver, pp.319-325, 1981.

L.GRANQUIST, On the Role of Editing, Statistical Review, pp. 105-118, 1984.

G.R.HAYES, Microcomputers and Minipopulations: The 1981 Cook Islands Census, in Asian and Pacific Census Forum, November, 1983.

F. HAYES-ROTH, D.A.WATERMAN, D.B.LENAT, Building Expert Systems, Addison-Wesley, 1983.

DM.HAWKINS, The detection of errors in multivariate data using principal components, J Amer Statist Ass, 69, pp.340-344, 1974.

B.JONES, Computer assisted design of experiments. PhD thesis, University of Kent at Canterbury, 1976.

B.JONES, The computer as a statistical consultant, BIAS, 7(2), pp.168-195, 1980.

J.B.KADANE, Some Statistical Problems in Merging Data Files, Compendium of Tax Research, pp. 159-179, Washington D.C., US Government Printing Office, 1978.

N.D.KODIKARA, S.ABEYASEKERA, INSTAT - A statistics package, In the proceedings of the Fifth National Computer Seminar organised by the Computer Society of Sri Lanka, 1975.

P.KORZENIOWS II, Tools of the Trade, in Computerworld, December 31, 1984/January 7, 1985.

M.KRAMER, Faster, Reliable Modems Changing Rules of the Road, in PC Week, December 25, 1984.

C.A.KULI DWS II, Artificial intelligence methods and systems for medical consultation, IEEE Trans. Patt. Anal. Machine Intell., PAMI-2, pp.464-475, 1980.

M.K.LEDDY, B.FREEDMAN, Storage Technology Takes Leap Forward in '84, in PC Week, December 25, 1984.

L.LYBERG, Control of the Coding Operation in Statistical Investigations: Some Contributions, Urval, No. 13, Statistics Sweden, Ph.D. thesis, 1971.

V.MANUKULASURIYA, N.D.KODIKARA, A Communication link between the BBC microcomputer and the Data General Minicomputer, 1985.

A.MARTIN, The challenge to market statistical products: strategies for expanding recovery of program costs.

G.MINTON, Inspection and correction error in data processing, J Amer Statist Ass, 64(328), pp.1256-1275, 1969.

G.MINTON, Verification error in single sampling inspection plans, J Amer Statist Ass, 67(337), pp.46-54, 1972.

L.NORDBÄCK, The CONDUCTOR, Workshop on the Statistical Computing Project, Budapest, Hungary, 12-17 November 1984, SCP/DE/WP89.

L.OLSSON, B.SUNDGREN, Towards a Long-Range Plan for the Restructuring of Statistical Production Systems in Sweden, Seminar on Integrated Statistical Information Systems, Bratislava, Czechoslovakia, 10-14 September 1979, CES/S M.11/R.2.

N.CURUS OFF, Information Processing Technology Planning in Malawi: The National Statistical Office and Ministry of Agriculture, 1983.

R.PIKE, The Blit: A Multiplexed Graphics Terminal, AT&T Bell Laboratories Technical Journal, Vol.63, No.8, Part 2, 1984.

D.B.RADNER, Construction of the 1972 Statistical Match File: Part I Statistical Matching Steps, Studies in Income Distribution, No.13, Social Security Administration, 1979.

J.L.ROOS, Artificial Intelligence and Macro-Economy, Eighth International Conference on Artificial Intelligence, Karlsruhe, Proceedings W Kaukman ed, Los Altos California, 1983.

N.RUGGLES, R.RUGGLES, E.WOLFF, Merging Microdata: Rationale, Practice and Testing, Annals of Economic and Social Measurement, Vol. 6, pp.429-444, 1977.

G.SANDE, Numerical edit and imputation, International Association for Statistical Computing, 42nd Session of the International Statistical Institute, 1979.

A.M.R.S MITH, L.S.LEE, D.J.HAND, Interactive user-friendly interfaces to statistical software, The Computer Journal, 26, pp.199-204, 1983.

K.SULLIVAN, Flash in the Pan, in Computerworld, December 31, 1984/January 7, 1985.

D.J.SPIEGELHALTER, Handling uncertainty in medical expert systems.

SUN, User's Manual for The Sun Workstation, Revision E, Sun Microsystems, Inc, 1984.

B.SUNDGREN, Statistical Data Processing Systems: Architectures and Design Methodologies, Proceedings of the Indian Statistical Institute Golden Jubilee International Conference, Calcutta, 16-19 December 1981, pp. 548-578.

B.SUNDGREN, Conceptual Design of Data Bases and Information Systems, Lecture Notes, University of Linköping and Statistics Sweden, 1984.

S.TAGG, The user interface of the data analysis package: some lines of development, Int J Man Machine Studies, 14, pp.297-314, 1981.

U.S. Office of Management and Budget, Guidelines for the Conduct of Matching Programs, Federal Register, March 30, 1979.

Numeri pubblicati nella collana "Quaderni di discussione":

- 84.01 REY G.M.
Le statistiche ufficiali e l'attività
della Pubblica Amministrazione
Giugno 1984
- 85.01 CRESCENZI F.
Nota su alcune metodologie per la
classificazione di unità territoriali
Febbraio 1985
- 85.02 CORTESE A.
Alcune considerazioni sulle prospettive
del censimento della popolazione
Marzo 1985
- 85.03 MATURANI G.
Stima delle ore di lavoro effettivamente
prestate dai lavoratori occupati
negli anni 1960-1983
Aprile 1985
- 85.04 NAPOLITANO P.
Esposizione di alcune tecniche per la
investigazione dei dati
Maggio 1985
- 85.05 RUSSO A.
Su un metodo di stima degli effetti
stratificazione e clustering e dell'ef-
fetto complessivo del disegno di cam-
pionamento nei campioni a due stadi con
stratificazione delle unità di primo
stadio
Settembre 1985

- 85.06 RUSSO A. e FALORSI P.
Rilevazioni campionarie delle forze di lavoro. Metodologia del campionamento calcolo e presentazione errori campionari
Novembre 1985
- 85.07 PAGNANELLI F.
Natimortalità, mortalità perinatale, mortalità infantile nel Comune di Napoli negli anni dal 1976 al 1980
Dicembre 1985
- 85.08 STEFANUTTI DE SIMONE L.
Le componenti stagionali delle variazioni dei prezzi al consumo dei prodotti alimentari
Dicembre 1985
- 86.01 RUSSO A.
Su un metodo di stima dell'effetto ponderazione nei campioni a due stadi con stratificazione delle unità primarie
Gennaio 1986
- 86.2 RUSSO A.
Una metodologia per la stima degli effetti stratificazione, clustering, ponderazione e dell'effetto complessivo nel disegno di campionamento nei campioni a due stadi con selezione delle unità primarie con reimmissione e probabilità variabile.
Gennaio 1986

- 87.01 DE NICOLA I. , CECCARELLI M., CALZARONI M.
Nota sulle statistiche nel settore della
edilizia e delle opere pubbliche
Gennaio 1987
- 87.02 MILITELLO A.
Un confronto tra redditi dichiarati al
Fisco e redditi stimati dalla Contabilità
Nazionale per gli anni 1981 e 1982
Febbraio 1987
- 87.03 MAROZZA F.
Centenario dell'International Statistical
Institute (ISI): adozione delle tecniche
informatiche per la statistica



