

EPI INFO

versione 6

EPI INFO

Versione 6

Manuale per il corso introduttivo

Programma (dischetti 1, 2, 3)

**Manuale d'uso completo in formato
Word 6 per Windows (dischetto 4)**

La Collana *Dossier* è pubblicata a cura di:
CDS (Centro di documentazione per la salute)
Azienda Usl Città di Bologna
via Triachini 17, 40138 Bologna, tel. 051/396311
Azienda Usl di Ravenna
via De Gasperi 8, 48100 Ravenna, tel. 0544/409018

Regione Emilia-Romagna, Servizio prevenzione collettiva
via Aldo Moro 30, Bologna, tel. 051/283111

Copia del volume può essere richiesta al CDS o
al Servizio Epidemiologia e Statistica sanitaria, Azienda Usl di Ravenna

Redazione:

Maria Edoarda Fava, Dipartimento di Prevenzione, Azienda Usl di Ravenna

Stampa: Azienda Usl di Ravenna, marzo 1997.

EPI INFO versione 6

Progetto del programma di: Andrew G. Dean, Jeffrey A. Dean, Denis Coulombier, Anthony H. Burton, Karl A. Brendel, Donald C. Smith, Richard C. Dicker, Kevin M. Sullivan e Robert F. Fagan

Programma di: Jeffrey A. Dean, Denis Coulombier, Donald C. Smith, Karl A. Brendel, Thomas G. Arner e Andrew G. Dean

Manuale originale di: Andrew G. Dean, aprile 1994

Questi programmi per computer IBM-compatibili sono stati prodotti grazie alla collaborazione tra:

Division of Surveillance and Epidemiology
Epidemiology Program Office

Centers for Disease Control and Prevention (CDC), Atlanta, Georgia 30333
Global Programme on AIDS Organizzazione Mondiale della Sanità Ginevra, Svizzera

Il manuale e i programmi sono di dominio pubblico e possono essere liberamente copiati, tradotti e distribuiti senza restrizione alcuna.

Questa versione italiana è stata curata da:

Marco Biocca, Laura Ciccolallo e Claudia Galassi
Centro di documentazione per la salute, Aziende USL Città di Bologna e Ravenna

Pasquale Falasca
Direzione sanitaria, Azienda Usl di Ravenna

Giovanni Falasca
Consulente software

Fabio Ostanello
Istituto di patologia e igiene veterinaria, Università di Padova

Francesco Saverio Violante
Servizio di medicina del lavoro, Azienda ospedaliera S. Orsola Malpighi, Bologna

La traduzione italiana del manuale della versione 5 (su cui si basa una parte del presente Manuale) è stata curata da:

Associazione ricerca e ambiente (AREA) (G. Gherardi, G. Giacomozzi, V. Lodi, G. Mancini, S. Mattioli, M. Missere, F.S. Violante) per incarico del SEDI, Usl 28 Bologna

La parte di programma in italiano è stata tradotta da:
Giovanni e Pasquale Falasca, che hanno anche curato
il Manuale per il corso introduttivo

Già da alcuni anni la Regione Emilia-Romagna, in collaborazione con l'Istituto superiore di sanità, ha diffuso attivamente in Italia il software Epi Info, prodigandosi per promuoverne la circolazione e la adozione fra gli operatori sanitari.

Epi Info è un ottimo ausilio per le ricerche epidemiologiche, e questo per due buone ragioni. Innanzitutto l'Organizzazione mondiale della sanità ha adottato questo tipo di software per uniformare l'introduzione dell'epidemiologia nei paesi in via di sviluppo ed ha contribuito in prima persona alla sua realizzazione. In secondo luogo il valore di Epi Info risiede in alcune caratteristiche peculiari al programma stesso: la varietà degli strumenti disponibili, l'ampiezza delle applicazioni e, non ultima, l'estrema facilità d'uso.

Se da una parte Epi Info mette a disposizione strumenti statistici sofisticati - all'altezza, per intendersi, dei sistemi SAS e SPSS ampiamente in uso fra gli specialisti - dall'altra è accessibile anche all'utente non specialista, che in poco tempo è in grado di apprezzarne la flessibilità agli usi sul campo più svariati. Mediante l'apprendimento pratico del programma su computer è possibile, inoltre, introdurre i non specialisti alle tecniche e alle problematiche degli studi epidemiologici. Non è esagerato affermare, allora, che Epi Info può anche essere uno strumento utile per la diffusione delle tecniche epidemiologiche fra gli operatori sanitari e, aggiungerei, per sensibilizzare a una visione più generale dei problemi sanitari.

Date queste premesse, non si può che accogliere favorevolmente questa traduzione italiana della documentazione a Epi Info curata da una molteplicità di soggetti. In particolare hanno collaborato - oltre agli operatori dell'Azienda USL di Ravenna - il Centro di documentazione per la Salute, l'Agenzia sanitaria regionale e l'Azienda ospedaliera S.Orsola-Malpighi.

Finalmente viene eliminata la barriera linguistica che rendeva troppo ristretta l'adozione di Epi Info in Italia, e possiamo considerare questo lavoro come un servizio reso all'avanzamento stesso delle ricerche epidemiologiche. Ci auguriamo dunque un'ampia diffusione di questo prodotto, che non potrà che giovare alle carenze del Sistema informativo sanitario italiano.

*Alessandro Martignani
Direttore Generale
Azienda USL di Ravenna*

*Mario Zanetti
Direttore Generale
Agenzia sanitaria regionale*

NOTE INTRODUTTIVE

La prima tappa di un'introduzione di Epi Info in Italia è stata la traduzione, nel 1992, del Manuale d'uso di Epi Info versione 5 e del Manuale per il corso introduttivo. Il volume pubblicato nella collana Dossier (n. 12) della Regione Emilia-Romagna conteneva le istruzioni essenziali all'uso di Epi Info, e includeva due dischetti contenenti il programma originale in inglese e la traduzione italiana del manuale d'uso completo.

Successivamente sono uscite ulteriori versioni originali di Epi Info 6, che hanno introdotto parecchie novità importanti. Si è deciso, quindi, di affrontare la traduzione della nuova versione includendo anche il software, in modo che il prodotto fosse più facilmente utilizzabile. Questo impegno è risultato particolarmente impegnativo a causa del rapido susseguirsi degli aggiornamenti e va considerato una ulteriore tappa verso la pubblicazione di una vera e propria edizione italiana di Epi Info completa, sia nel software che nel Manuale d'uso.

Questo volume contiene il "Manuale per il corso introduttivo a Epi Info versione 6", che è una versione aggiornata rispetto al Dossier n. 12 pubblicato nel 1992. Fra le novità sono da citare l'introduzione di un nuovo capitolo su EPIGLUE (il software per creare programmi personalizzati), un capitolo sulle nozioni fondamentali sulle basi di dati, e un capitolo contenente una sintesi di rapida consultazione dei comandi di EPED e ANALYSIS.

Acclusi al manuale ci sono quattro dischetti contenenti, i primi tre, i file di installazione di Epi Info, e il quarto una copia del Manuale d'uso completo in italiano già incluso nel software, ma in formato Word per Windows 6, in modo da poter essere stampato liberamente. Nel software di Epi Info 6 si troveranno tutti i file di esempio, i file didattici e l'interfaccia utente, interamente tradotti in italiano. Inoltre, dal menu Manuale, è consultabile in linea il Manuale elettronico in forma ipertestuale.

Il software che viene distribuito in questa occasione è l'ultimo disponibile, la versione 6.04 dell'ottobre 1996. Il manuale d'uso è invece quello della versione 6.0 ma, nel menu Manuale, è stata inserita una voce sulle novità della versione 6.04. Il programma EPI6.EXE (il menù principale di Epi Info) è stato tradotto integralmente e ricompilato, utilizzando i programmi sorgente originali, così come il programma di installazione, che ora è integralmente in italiano. La traduzione integrale del software dai file sorgente potrà essere messa a disposizione nel prossimo futuro, grazie alla modularità e alla flessibilità di Epi Info. Nel frattempo non si arresterà la diffusione di Epi Info e si potrà fare pratica sulla versione 6.

È possibile, ovviamente, che questo prodotto contenga degli errori e la collaborazione nel segnalarceli sarà particolarmente gradita.

Non ci resta dunque che rimandare alla lettura dei capitoli 1 e 2 per conoscere le novità della versione 6, e al capitolo 3 per l'installazione del software in DOS (e, novità, in Windows 3.1 e Windows 95) e dare appuntamento alla versione italiana integrale di Epi Info.

Ravenna, 20 febbraio 1997

INDICE

<i>Capitolo 1.</i>	Introduzione al corso	11
<i>Capitolo 2.</i>	I programmi di Epi Info 6	17
<i>Capitolo 3.</i>	Installazione e avvio di Epi Info	21
<i>Capitolo 4.</i>	Creare questionari usando l'editore EPED	25
<i>Capitolo 5.</i>	Cosa bisogna sapere sulle basi di dati.....	31
<i>Capitolo 6.</i>	Inserire i dati usando il programma ENTER.....	41
<i>Capitolo 7.</i>	Il programma CHECK: controllo degli errori e dei codici.....	49
<i>Capitolo 8.</i>	ANALYSIS: liste, frequenze, tabelle, statistiche e grafici.....	57
<i>Capitolo 9.</i>	Scrivere programmi e impostare i dati per ANALYSIS.....	67
<i>Capitolo 10.</i>	Ancora su ANALYSIS	81
<i>Capitolo 11.</i>	STATCALC ed EPITABLE: calcolatori statistici	91
<i>Capitolo 12.</i>	VALIDATE: doppia immissione di dati e verifica	95
<i>Capitolo 13.</i>	EXPORT: produrre file per altri sistemi.....	97
<i>Capitolo 14.</i>	Importare dati da altri programmi	101
<i>Capitolo 15.</i>	EPIGLUE: un sistema informativo epidemiologico e di menu	105
<i>Capitolo 16.</i>	Sintesi dei comandi di Epi Info.....	113
<i>Capitolo 17.</i>	Glossario breve dei termini epidemiologici fondamentali.....	139
	Indice degli argomenti	145

Capitolo 1

INTRODUZIONE AL CORSO

“La vera gerarchia dei fenomeni, come la vera gerarchia dei rapporti, si incarna e prende forma su un piano totalmente diverso da quello delle regole convenzionali.”

Igor Stravinsky

Questo corso introduttivo a Epi Info permette di acquisire alcune abilità all'uso di base del personal computer, introduce a particolari tecniche e procedure avanzate per la raccolta e l'archiviazione dei dati, nonché ad acquisire abilità nel campo della elaborazione e nello scambio sistematico delle informazioni, necessarie sempre più spesso per svolgere correttamente le complesse attività operative nel settore della sanità. Va sottolineata l'importanza strategica che assume Epi Info per lo sviluppo del sistema informativo e per la sua integrazione in ambito sanitario.

Le sue caratteristiche permettono la estrema disseminazione dell'informazione e la sua fruizione, rendendola disponibile sia ai livelli decisionali e gestionali, sia all'esercizio operativo delle prestazioni. Epi Info permette la piena accessibilità delle informazioni da parte dei soggetti coinvolti nel sistema informativo sanitario: operatori sanitari, amministratori e utenza - un concetto che in altri tempi veniva espresso come “concretezza della partecipazione” (1) -. Tale accessibilità viene intesa in accordo con la definizione di Donabedian: “L'accessibilità deve essere considerata qualcosa oltre la mera presenza o 'disponibilità' di risorse in un dato momento e posto. Essa include le caratteristiche delle risorse che facilitano o offrono la modalità di utilizzo da parte dei potenziali clienti” (2).

Non è possibile, oggi, elaborare o analizzare dati sanitari senza l'aiuto di un personal computer. Qualche anno fa si rifiutava il calcolatore elettronico per il timore (o il rischio) della eccessiva centralizzazione dell'informazione, si temeva di privarla della sua ricchezza con processi di formalizzazione difficili da controllare. (3 4).

In biostatistica il computer ha facilitato i pesanti e lunghi calcoli, ha aperto vie a nuove direzioni di ricerca.

La micro-informatica ha, inoltre, messo a disposizione questi strumenti alla portata di tutti: oggi i personal computer sono molto diffusi nei servizi sanitari.

L'importanza dell'informatica come strumento di lavoro nel settore sanitario è incontestabile (5 6).

1 F. Bottasso, M. Desideri, P. Santacroce, *Modello informativo del sistema*, in "La politica dell'informazione nel sistema socio-sanitario" Regione Emilia-Romagna, Toscana, Umbria, Villalago 1977

2 A. Donabedian, *Aspect of medical care administration*, Harvard University Press, 1973

3 F. De Rosis, *Informatica sanitaria: mercato artificiale o strumento di partecipazione?*, "Epidemiologia e Prevenzione" n.1, p. 8-16; 1977

4 Regione Emilia-Romagna, Toscana, Umbria, *Documento sulla informatica sanitaria*, 1977

5 Oms, *Informatique, Télématique et Santé, Applications actuelles et potentielles*, Genève, 1990

6 W.W. Christensen, E. Stearns, *Microcomputer in Health Care Management*, Aspen Publishers Inc., 1990

D'altra parte l'utilizzo di questi apparecchi, per la ricerca epidemiologica, è stata condizionata da programmi pesanti, costosi e difficili da imparare ad usare come l'SPSS, il SAS, il BMDP, il SYSTAT, ecc.

Alcuni ambienti specializzati hanno studiato la possibilità di agevolare le applicazioni nel campo dell'epidemiologia e hanno sviluppato una intera generazione di programmi che "girano" su personal computer (7 8 9 10) che permettono di avere un interesse generale, non solo accademico, per praticare la ricerca e per gestire la propria informazione di "esercizio" dell'attività di produzione di servizi a livello locale. Epi Info è uno di questi programmi che permette di praticare (e insegnare) l'epidemiologia "de poche" (da taschino, come dicono i francesi) (11).

Alla rivoluzione della messa in opera di metodi legati all'automazione ne sta seguendo un'altra relativa al loro apprendimento. L'insegnamento assistito dall'elaboratore conosce progressi che meritano molta attenzione. "Gli inconvenienti dell'insegnamento magistrale sono conosciuti: focalizzano soprattutto sul meglio o le cose meno buone, o le più frequenti.... non riesce ad adattarsi in ogni momento al livello e alla qualità di ciascuno. Un altro limite dell'insegnamento classico è la passività.... Ci vorrebbe un insegnamento adattato a ciascuno" (12).

La tecnica didattica *dell'apprendimento attivo* cerca di saldare il momento dell'azione svolta nell'attività quotidiana con quello dell'apprendimento (Dewey). Ha successo quando il discente ha un "motivo" per imparare; quando avverte che l'oggetto di studio supplirà a una mancanza che produce un qualche conflitto nella sua situazione attuale. (13)

In ultima analisi, se solo in virtù di una necessità pratica siamo costretti a discriminare le cose collocandole in categorie puramente convenzionali (primarie - secondarie, principali - subordinate) il nostro obiettivo non sarà quello di separare gli elementi di cui ci occuperemo, ma di distinguerli senza separarli, osservandoli nel loro contesto naturale. Si adotterà una sorta di parallelismo, un metodo di "sincronizzazione", vale a dire che si uniranno i principi generali ai fatti concreti e particolari mediante argomentazioni che servono a questi ultimi per arrivare ai primi.

7 D.Donabedian, *Computer-Taught Epidemiology*, The American Journal of Nursing, vol 14, n.12, p.749-751, 1976

8 F. De Rosis, S.Pizzullo, *Epistat*, "Epidemiologia e Prevenzione" n.28, p.54-56; 1986

9 S.Pizzullo, *Analizzare dati in epidemiologia con un personal computer: problemi ed integrazione fra software diversi*, "Epidemiologia e Prevenzione" n.30 p.40-43; 1987

10 G.Rodriguez *et al.*, *A computer program for application in epidemiological analysis*, "European Journal of Epidemiology", vol 9, n.1 p. 1-4, 1993

11 Epi-Info Version 5, Traduction française par R.Freund, B.Junod et H.Saout de l'Ecole Nationale de Santé Publique. Editions ENSP, Reimes 1991.

12 D. Schwartz nell'introduzione a: AJ. Valleron, P. Lazar, *Exercices programmés de statistique à l'usage des médecins et des biologistes*, Paris : Flammarion Médecine-Sciences, 1967

13 Mentre non si crea il bisogno di alcuno sforzo se tutto "va liscio", collegare il progetto educativo ai problemi concreti di lavoro dei partecipanti, sostituendo il più possibile le esercitazioni con i casi reali, stimolando i partecipanti del corso ad una rielaborazione delle soluzioni proposte e ad una scoperta originale per risolvere i propri problemi (via praticabile con l'insegnamento assistito da computer) è una tecnica didattica che si richiama a modelli di apprendimento di tipo pragmatico, all'esperienza diretta di chi chiede soluzioni a strumenti e tecniche conoscitive (come Epi Info o l'epidemiologia), e la produzione di un effettivo miglioramento cognitivo.

CHE COS'É EPI INFO

Epi Info contiene una serie di strumenti per microcomputer per:

- raccogliere dati epidemiologici con un questionario,
- creare e gestire i piccoli archivi in modo rapido e maneggevole,
- organizzare il disegno di uno studio,
- produrre indicatori epidemiologici in modo relativamente facile,
- elaborare ogni tipo di base dati in proprio,
- scrivere i risultati delle elaborazioni in un rapporto,
- gestire in proprio un sistema di sorveglianza.

Le caratteristiche che distinguono Epi Info sono così riassunte:

- semplicità per un uso quotidiano,
- quattro giorni di corso per acquisire le abilità fondamentali,
- copre il 40% dei comandi dei grossi programmi, permettendo di rispondere all'80% dei bisogni quotidiani di analisi dei dati,
- i professionisti possono condividere l'uso del programma conservando l'attenzione agli obiettivi del proprio lavoro,
- produzione immediata dei risultati: non necessita del ricorso all'informatica o il doversi specializzare in altre discipline,
- gratuito: si adatta a tutti i tipi di Personal Computer,
- epidemiologia *de poche*: permette una epidemiologia diffusa e gestisce facilmente informazioni per la presa di decisione a tutti i livelli di gestione. Inoltre è adatto allo sviluppo di tecniche di comunicazione dell'informazione in prossimità degli utenti finali.

OBIETTIVO GENERALE DEL CORSO

Il corso si prefigge di fornire le conoscenze necessarie ad adoperare lo strumento Epi Info come un **utensile** per risolvere problemi conoscitivi, determinare la raccolta e la archiviazione ordinata dei dati, per la elaborazione di informazioni sintetiche utili a migliorare la qualità del proprio lavoro.

OBIETTIVI SPECIFICI

- Conoscere gli elementi costitutivi di una base di dati.
- Saper progettare e costruire un questionario usando EPED

- Creare una base di dati in Epi Info e caricare i dati.
- Creare un file CHECK per controllare il corretto inserimento dei dati.
- Usare ANALYSIS per produrre rapporti ed elaborare dati.
- Incorporare in una sola base di dati archivi di differente formato.
- Importare in Epi Info dati da vari formati.
- Esportare le basi di dati di Epi Info in altri formati.
- Imparare come si accede alla documentazione di Epi Info.

L'apprendimento di Epi Info è strutturato in tre ambiti di esperienza, ciascuno legato all'apprendimento di differenti abilità, distinte in tre livelli:

- I. CONOSCERE EPI INFO (esperienza sulle potenzialità del sistema)
- II. USARE EPI INFO (adoperare applicazioni predefinite a specifici settori)
- III. PRATICARE EPI INFO (raccogliere dati, ricercarli, tabularli, graficarli, calcolare tassi...).

Abilità di I livello

- Avviare il menù principale
- Scrivere il questionario in EPED
- Introdurre i dati nel questionario usando ENTER
- Inserire il controllo degli errori, salti e codifica automatica usando CHECK.
- Analizzare i dati introdotti usando ANALYSIS per produrre liste, frequenze, tavole a più entrate, medie, grafici e alcune statistiche associate.

Abilità di II livello

- Controllare gli errori di inserimento, saltare dei campi a seconda delle informazioni fornite e inserire delle codifiche automatiche nel questionario, utilizzando il programma CHECK.
- Selezionare record, creare nuove variabili, ricodificare i dati, manipolare le date e inserire delle condizioni mediante l'operatore IF, all'interno di ANALYSIS.
- Incorporare queste operazioni in file di programma così da poter essere eseguite ripetutamente, anche da altre persone inesperte del programma.
- Importare ed esportare file da altri sistemi come SAS, SPSS, dBASE e Lotus 1-2-3.
- Cambiare i nomi delle variabili nei file di dati utilizzando il programma CHECK.

Abilità di III Livello

Il terzo livello di applicazioni è importante se si sta progettando un sistema permanente di database, un vasto studio o se si vogliono usare le funzioni di Epi Info per scopi particolari. È possibile:

- Programmare il processo di inserimento dati includendo operazioni matematiche, controlli logici, cambio di colori, finestre *popup* e routine personalizzate scritte con altri linguaggi di programmazione.
- Specificare il formato dei report da ANALYSIS in modo da produrre tabulati standard.
- Inserire dati in più di un file durante la stessa sessione, spostandosi automaticamente attraverso i vari questionari con ENTER., inserire messaggi di aiuto e altre facilitazioni all'introduzione dei dati.
- Mettere in relazione diversi tipi di file in ANALYSIS così da poter ottenere dei risultati che richiedono dati da più di un file (analisi relazionale).
- Confrontare due file duplicati inseriti da differenti operatori al fine di individuare errori di inserimento dati.
- Usare i calcolatori statistici (STATCALC, EPITABLE) ed effettuare analisi dei principali indicatori epidemiologici calcolati.
- Progettare e realizzare un sistema di reporting ipertestuale con EPIGLUE.

Il corso si adatta alle caratteristiche di conoscenza già acquisita dai discenti o in funzione dell'approfondimento desiderato articolando l'insegnamento sui tre livelli di abilità descritti. La particolare flessibilità del software consente di orientare il corso verso **INSEGNAMENTI FUNZIONALI**, cioè a strutture didattiche maggiormente indirizzate alla acquisizione delle abilità necessarie alla **soluzione di specifici problemi conosciuti**.

Il corso contenuto in questo manuale copre il primo livello per intero, parte del secondo e alcuni cenni sul terzo. Per coloro che volessero approfondire il secondo e il terzo livello, si potrà utilizzare il manuale elettronico in linea (menu Manuale) contenuto nella finestra principale di Epi Info. Si tratta di un manuale ipertestuale, interamente tradotto in italiano, con la stessa suddivisione in capitoli e paragrafi dell'edizione del manuale cartaceo inglese. Mediante gli indici ipertestuali, evidenziati in giallo, sarà semplice, dall'indice del capitolo, arrivare rapidamente all'argomento interessato.

Capitolo 2

I PROGRAMMI DI EPI INFO 6

Epi Info è un pacchetto integrato di programmi per computer per manipolare dati epidemiologici in formato di questionario ed anche per organizzare progetti e risultati di studio in testi che possono formare parti di relazioni scritte. Un questionario può essere messo a punto ed elaborato in pochi minuti, ma Epi Info può anche costituire la base di un potente sistema di database per la sorveglianza sanitaria con molti file e tipi di record. Include le caratteristiche maggiormente utilizzate dagli epidemiologi nei programmi di statistica come SAS o SPSS, e programmi di database come dBASE, combinati in un unico sistema che può essere liberamente copiato e fornito ad amici e colleghi.

Come si è detto ci sono tre livelli di approccio in Epi Info per elaborare questionari o altri dati strutturati. Utilizzando Epi Info al livello più semplice, potete computerizzare un questionario o un modulo di indagine in pochi minuti:

- Avviando il menù principale.
- Scrivendo il questionario o il modulo con EPED (l'elaboratore di testi), o con un altro programma di scrittura.
- Inserendo dati nel questionario utilizzando il programma ENTER.
- Analizzando i dati col programma ANALYSIS per creare listati, tassi, tabelle di contingenza, medie, grafici e statistiche correlate.

Al crescere della conoscenza del programma, si potranno utilizzare caratteristiche più avanzate per conformare l'inserimento e l'analisi dei dati alle proprie esigenze.

EPED, l'elaboratore di testi incluso nella Versione 6, contiene un sistema di programmazione di testi denominato EPIAID che consente di utilizzare tutte le caratteristiche di Epi Info per strutturare i dati. Sono inclusi programmi (tradotti in italiano), scritti in EPIAID, che guidano attraverso la creazione di questionari e la progettazione di un'indagine epidemiologica. Il testo originato può poi essere utilizzato come parte di una relazione di indagine. Il programma EPIAID può essere utile anche per la scrittura di lettere, la stesura di relazioni o per altri scopi di videoscrittura.

La Versione 6 di Epi Info contiene una raccolta di nuove caratteristiche, come il sistema di menu a tendina completamente configurabile, nuovi comandi per programmare l'inserimento e l'analisi dei dati, la capacità di ordinare e mettere in relazione file molto estesi, un programma per analizzare i dati di un'indagine campione complessa, un nuovo calcolatore epidemiologico, e un programma per l'antropometria nutrizionale. I file creati con le precedenti versioni di Epi Info possono essere usati nella Versione 6. Con poche revisioni, anche i programmi scritti per la Versione 5 possono essere avviati nella Versione 6.

I programmi inclusi in Epi Info sono:

EPI6 - Il menù principale che permette l'accesso agli altri programmi.

EPED - Un elaboratore di testi per creare questionari specifici o di utilizzo generale. Uno speciale strumento ausiliario chiamato EPIAID contenuto in EPED avvia programmi che assistono nella preparazione di testi e nella progettazione di indagini epidemiologiche.

ENTER - Produce automaticamente un file di base di dati a partire da un questionario creato con EPED o un altro programma di scrittura. Permette l'inserimento di dati da un questionario in un file su disco e consente la revisione del formato del file anche dopo l'inserimento dei record.

ANALYSIS - Produce elenchi, tabelle di contingenza e una varietà di altre elaborazioni da file di Epi Info o di dBASE. Accompagnano tali tabelle alcune statistiche epidemiologiche corrette, quali: stima del rischio relativo, odds ratio, limiti di confidenza, il test esatto di Fisher e i test chi-quadrato. Sono inoltre a disposizione analisi stratificate con la tecnica di Mantel-Haenszel, i test unidirezionali ANOVA e Kruskal-Wallis, la regressione lineare e l'analisi degli studi caso-controllo accoppiati. I record possono essere selezionati o estrapolati utilizzando variabili definite, iterazioni "if" ed operazioni logiche e matematiche. Sono inclusi in ANALYSIS grafica, formattazione di relazioni complesse, generazione di nuovi insiemi di dati da saggiare ed un linguaggio di programmazione. Le variabili sparse in diversi file possono essere congiunte ed analizzate come su una singola unità.

CHECK - Controlla i valori consentiti; gestisce elenchi di valori, codifiche automatiche e salti di campo per il programma ENTER. È possibile inoltre, grazie a CHECK, compiere operazioni matematiche e logiche tra campi, saltare da un campo a un altro a seconda dei valori che si stanno inserendo, accedere a diversi file nella stessa maschera di immissione dati e, infine, utilizzare in CHECK programmi esterni forniti dall'utente.

CSAMPLE - Compie analisi di dati che prevedono strategie di campionamento complesse, prendendo in considerazione l'analisi per cluster, la stratificazione e il peso.

STATCALC - Calcola statistiche da valori tabellari inseriti dalla tastiera. STATCALC permette calcoli da tabelle due per due, singole e stratificate; il dimensionamento del campione e l'analisi delle tendenze singole e stratificate.

EPITABLE - Un calcolatore epidemiologico, consente di calcolare molti indicatori epidemiologici di uso corrente.

EPINUT - Un programma di antropometria nutrizionale provvisto di una procedura per il calcolo di indici antropometrici.

EXPORT - Esporta i file di dati da EPI INFO in 12 formati diversi per programmi di database e di statistica.

IMPORT - Importa i file di altri programmi in modo da poter essere utilizzati in EPI INFO.

MERGE - Unisce file di dati prodotti a partire da questionari con lo stesso o con altri formati. Ciò permette la combinazione di file di dati inseriti in computer diversi e l'aggiornamento di precedenti record servendosi dei dati provenienti da nuove registrazioni.

VALIDATE - Confronta due file di Epi Info prodotti da diversi operatori e ne riporta ogni differenza.

Manuale elettronico in italiano - Il manuale elettronico in linea contenuto nella finestra principale di Epi Info (menu Manuale), è un ipertesto, interamente tradotto in italiano, con la stessa suddivisione in capitoli e paragrafi dell'edizione del manuale cartaceo inglese. Mediante gli indici ipertestuali, evidenziati in giallo, sarà semplice, dall'indice del capitolo, arrivare rapidamente all'argomento interessato.

Didattica - Nella sezione didattica, del menu principale, sono contenuti quattro programmi interattivi, tradotti in italiano, che aiutano l'auto-apprendimento delle principali caratteristiche di EPED, ENTER e ANALYSIS.

Esempi - A titolo di esempio sono forniti: due indagini epidemiologiche e un sistema di calcolo antropometrico in indagini nutrizionali. Sono inoltre forniti tre esempi di indagini campione complesse da utilizzare con CSAMPLE; infine ci sono numerosi altri piccoli file di dati e altri esempi sono forniti nel manuale.

Epi Info, Versione 6, richiede un microcomputer IBM-compatibile fornito del sistema operativo PC-DOS o MS-DOS (Versione 2.0 o successive), 512 K (524,288) byte di memoria RAM e almeno un drive per floppy disk. È richiesta inoltre una scheda grafica per riprodurre grafici: sono supportate le schede Hercules o IBM monocromatiche, EGA, VGA, IBM3270, IBM8514 e AT&T. Per la stampa, vengono forniti i driver per 30 diversi tipi di stampanti. Per un utilizzo efficace di Epi Info, quando si utilizzano files di grandi dimensioni, è consigliato un hard-disk, 640 K di memoria e scheda grafica e monitor entrambi a colori.

Se nel computer è installato un gestore della memoria (EMM386 o QEMM), Epi Info riesce a utilizzare limitatamente la RAM al di sopra dei 640 K. I programmi con un overlay (.OVR) saranno caricati nella memoria alta e gireranno più velocemente; ma non necessariamente si libererà più memoria convenzionale per il programma.

I file di dati possono contenere tanti record quanti ne permettono il DOS e la capacità del disco rigido (fino a due miliardi). Un questionario può avere fino a 500 linee, approssimativamente 20 schermate. Il numero delle variabili non ha limiti, a patto che si adattino entro le 500 linee. La lunghezza massima di una variabile di testo è di 80 caratteri. La lunghezza totale delle variabili in un file non deve eccedere i 2048 caratteri. Data la possibilità di legare diversi file correlati, ciò non costituirà un problema pratico, dal momento che si possono distribuire i dati in diversi file, e metterli in relazione fra loro durante l'analisi. I programmi in Epi Info richiedono circa 6 megabyte di spazio su disco, ma ANALYSIS, il programma più esteso, riesce a stare in un dischetto da 360K.

La Versione 5 di Epi Info è stata prodotta dalla CDC e dalla Global Programme on AIDS dell'OMS. La versione 6 invece è in gran parte opera degli autori del CDC.

I programmi sono resi disponibili dall'OMS e dal CDC e non c'è copyright. È consentito e raccomandato farne copie per altri e distribuirle. Si prega di fornire copie di questo manuale quando si distribuiscono i programmi. Se si progetta una distribuzione regolare e sistematica del programma, è opportuno avvisare gli autori sul numero approssimativo di copie distribuite, dal momento che la CDC tiene un registro delle copie del programma distribuite nel mondo. Si consiglia di testare le copie del programma con un software antivirus ufficialmente riconosciuto.

Capitolo 3

INSTALLAZIONE E AVVIO DI EPI INFO

INSTALLAZIONE

I programmi nei dischi di distribuzione Epi Info sono nel formato compresso e debbono essere espansi prima che il programma sia avviato. Un programma chiamato INSTALLA è contenuto nel disco 1 e questo programma deve essere usato per installare il sistema.

Per installare Epi Info, accendere il computer nel modo usuale fino a che non compare il messaggio di richiesta del sistema operativo.

C:\>

Posizionare il disco 1 del sistema di Epi Info nel drive A (o nel drive B) e digitare:

A:INSTALLA se si è inserito il disco nel drive A

B:INSTALLA se si è inserito il disco nel drive B

A questo punto sarà sufficiente seguire le istruzioni a video per portare a termine l'installazione.

Il programma richiederà le seguenti informazioni:

- quale è il drive SORGENTE, quello nel quale è inserito il primo disco di installazione. Usualmente sarà il drive A e si dovrà digitare la lettera A (senza i due punti).
- Quale è il drive di DESTINAZIONE in cui installare o copiare i file. Se si sta installando il sistema su un disco rigido, questa sarà la lettera corrispondente all'hard disk, normalmente C. Per installare Epi Inf su floppy disk (si ricordi che nessun programma di Epi Info è di dimensioni maggiori di 360 Kb, pertanto perfino ANALISYS può essere eseguito su un vecchio floppy da 5" 1/4) digitare la lettera di un altro drive di floppy disk, di solito B.
- Questo drive (DESTINAZIONE) è un disco rimovibile (un floppy) rimovibili (Y/N)? Si risponde "Y" se il disco di destinazione è un floppy disk, altrimenti "N".
- Si vuole installare il sistema per l'uso, o copiarlo per ulteriore distribuzione (selezionare "I" o "C")? Si deve rispondere "I" per la normale installazione su disco rigido. Se si risponde "C", verranno semplicemente riprodotti altri dischi floppy di installazione di Epi Info. Che tipo di scheda video è installata nel computer che userà il sistema? Sono presenti le opzioni per le schede CGA, MCGA, EGA o VGA, IBM8514, IBM 3270 PC, AT&T 6300, Hercules. Il programma è in grado di riconoscere da solo la scheda installata (inoltre, oggi, quasi tutti i computer montano schede VGA). Se comunque si hanno dei dubbi premere <F8> per selezionarle tutte.
- La Scelta dei Driver di stampa. I driver stampa sono file che aiutano Epi Info a stampare i grafici su stampanti diverse o a trasferire grafici a file formattati per altri sistemi. La scelta consigliata, se non si hanno problemi di spazio sul disco rigido (poche centinaia di Kb) è di selezionarli tutti. Altrimenti selezionare solo il driver per la propria stampante e i driver per i grafici a video. Dopo avere scelto i driver di stampa tramite i tasti freccia e la <barra spaziatrice>, premere <F4> per continuare.

- Selezione dei gruppi da installare. I file di Epi Info sono divisi in gruppi nei dischetti e possono esserne installati solo alcuni. Si consiglia di installare tutti i gruppi (l'intero sistema Epi Info non richiede più di 8-9 megabytes), si può comunque scegliere di escludere qualche gruppo, come l'esempio di sistema di sorveglianza NETSS. Per selezionare tutti i gruppi premere <F8>. Dopo aver selezionato i gruppi, premere <F4> per continuare.
- Il programma di installazione richiederà di inserire i dischi necessari con messaggi a video: premere <Invio> dopo aver inserito il disco richiesto. In ogni momento, eccetto durante la copia di un file, si può premere <F10> per interrompere l'installazione.

NOTA

Nel file **AUTOEXEC.BAT** (in genere presente nella directory principale C:\>):

Ci deve essere un comando di PATH che includa un percorso (path) per \EPI6 sul disco dove risiede il sistema di Epi Info. (Aggiungere nella riga del PATH "C:\EPI6;") I percorsi (path) precedenti debbono rimanere attivi.

Nel file **CONFIG.SYS**(in genere presente nella directory principale C:\>): ci dovrebbero essere i comandi: "FILES = 20" e "BUFFERS = 20". Questi numeri potrebbero essere superiori a 20, se il DOS o altri programmi hanno già inserito questi comandi.

Durante l'installazione, il programma tenterà di aggiornare questi file automaticamente previo avvertimento. Al messaggio di avvertimento, digitando "Y", i file preesistenti saranno rinominati AUTOEXEC.OLD e CONFIG.OLD. I nuovi AUTOEXEC.BAT e CONFIG.SYS saranno creati con le necessarie aggiunte, come descritto sopra.

MESSA A PUNTO DELLA STAMPANTE

Molti risultati di Epi Info possono essere stampati con qualsiasi stampante standard IBM- o Epson-compatibile. Se si possiede una stampante a matrice d'aghi, non è necessario alcun ulteriore adattamento per stampare i grafici prodotti con ANALYSIS. Se si possiede una stampante laser o un plotter, invece di impostare ogni volta ANALYSIS per la propria stampante, è preferibile creare un file, di nome CONFIG.EPI, contenente uno dei seguenti comandi:

SET PRINTER = EPSON (per stampanti a matrice d'aghi Epson/IBM -compatibili)

SET PRINTER = LQ1500 (per stampanti Epson LQ1500 compatibili)

SET PRINTER = HP (per stampanti laser Hewlett Packard compatibili)

SET PRINTER = PLOTTER (per plotter Hewlett Packard compatibili)

La maggior parte delle stampanti laser richiede pagine da 60 righe invece delle normali 66. Se si ha una stampante laser, si deve aggiungere al file CONFIG.EPI il comando PLINES = 60. Le stampanti connesse ad una porta seriale, anzichè parallela, richiedono inoltre comandi di PORT, BAUD RATE e PMODE come descritto nel Capitolo 36 del Manuale, nella sezione del comando SET.

Anche nel programma EPED si dovrà impostare, nel menù SET, la lunghezza della pagina a 60 righe, per stampanti laser, o a 66 per la maggior parte delle stampanti ad aghi. La lunghezza della pagina potrà essere selezionata in ciascuno dei tre modi in EPED:

TXT, WW, e QES. Ogni volta che si esce dal menu SET <F6>, se si sono apportate modifiche, ricordarsi di salvarle con il comando SAVE SETTINGS.

AVVIO DI EPI INFO

Per avviare Epi Info digitare dal prompt del DOS C:\>:

CD EPI6 <Invio>

EPI6

per entrare nella finestra del menu principale di Epi Info, dalla quale si potranno gestire tutti gli altri programmi. Se il programma non parte, probabilmente

Per questa edizione di Epi Info 6 si sono riscritti interamente i menu di Epi Info, per tradurli in italiano. Si è tradotta inoltre l'interfaccia di Epigluè in italiano e la relativa guida in linea, gli esempi del menu esempi e i menu File e Modifica con relative guide. Rimangono in inglese le guide in linea interne dei programmi di Epi Info attivate con <F1>, ma si è ovviato all'inconveniente con il manuale elettronico in italiano in formato ipertestuale (menu Manuale).

EPI INFO DA WINDOWS

Per questa edizione di Epi Info è prevista la possibilità di utilizzare il programma direttamente da Windows. Il vantaggio è tutto nella possibilità di utilizzare contemporaneamente più programmi (multitasking). Per esempio è possibile avviare ANALISYS e contemporaneamente lavorare su un questionario in EPED o cercare dei dati con ENTER, senza ogni volta uscire da un programma per entrare in un altro.

Basta ricordare che per passare da un programma all'altro, di quelli aperti in Windows, si utilizza una scorciatoia poco conosciuta ma comodissima: <ALT-TAB>. Si tiene premuto ALT e, ad ogni pressione del tasto TAB (il tasto per le tabulazioni), compare a video il nome di un programma attivo. Basta quindi rilasciare il tasto ALT appena appare il nome del programma con cui si vuole lavorare.

In fase di installazione, se sul computer è presente una copia di Windows, verrà automaticamente installato un file (Epi.grp) per creare una finestra aggiuntiva in Program Manager, contenente una serie di icone corrispondenti ai vari programmi contenuti in Epi Info. Sarà sufficiente cliccare su un'icona per entrare in Epi Info **senza uscire da Windows** e usufruendo pienamente di tutte le funzionalità multitasking.

Inoltre si noterà un'icona di Word per Windows, dal titolo "Manuale Epi Info 6.Doc". Si tratta della traduzione italiana del manuale originale di Epi Info 6. Il documento è nel formato di Word 6 per Windows ed è sufficiente cliccare sull'icona per aprire Word 6 e caricare automaticamente il documento dal dischetto 4 dell'edizione italiana di Epi Info. A questo punto si può scegliere di consultare il manuale a video, oppure di stamparlo. Naturalmente sul computer deve essere installata una copia regolare di Microsoft Word 6 per Windows.

CREAZIONE DEL GRUPPO: "PROGRAMMI DI EPI INFO" IN WINDOWS

Per adoperare Epi Info direttamente da Windows bisogna creare un nuovo gruppo di programmi in Program Manager. La procedura è molto semplice:

- Dal menu **File** di Program Manager cliccare l'opzione **Nuovo**
- Appare una finestra dialogo per scegliere fra **Programma** e **Gruppo di programmi**. Selezionare **Gruppo di programmi** e cliccare su **OK**.
- Appare un'altra finestra dove inserire un campo "descrizione" e un campo "file del gruppo". Digitare EPI.GRP nel campo "file del gruppo" e lasciare vuoto l'altro. Quindi cliccare su **OK**.
- Immediatamente apparirà una nuova finestra di Program Manager con i programmi di Epi Info.

Nota per i possessori di Windows95

In Windows95 non appare Program Manager ma, anche se non tutti lo sanno, è lo stesso presente nella directory di windows sotto il nome di Progman.exe. Per avere le icone di Epi Info in Windows95 si potrebbe così avviare Program Manager e seguire le stesse istruzioni di cui sopra.

Conviene però sfruttare il menu Avvio di Windows95 e creare un sottomenu per i programmi di Epi Info. Ecco come si procede:

- Cliccare su **Avvio** e poi su **Esegui**
- Nel campo del comando da eseguire digitare esattamente:
grpconv c:\windows\epi.grp
- Il sottomenu verrà creato immediatamente. Cliccare su **Avvio** nuovamente per controllare.

VISUALIZZARE E MODIFICARE FILE DAL MENU PRINCIPALE DI EPI6

Il menu EPI6 ha un editor incorporato che offre le funzioni di editing di base (scrittura e modifica) per i file di testo. Per creare un nuovo file, selezionate NUOVO dal menu FILE. Per aprire un file già esistente, selezionate APRI dallo stesso menu.

Consultare il capitolo 20 del manuale elettronico per ulteriori informazioni sui comandi dei menu FILE e MODIFICA.

Capitolo 4

CREARE QUESTIONARI USANDO L'EDITORE EPED

INTRODUZIONE

Prima di cominciare a inserire i dati, bisogna costruire un questionario con Epi Info. Fatto questo il programma ENTER costruirà automaticamente un file di dati, chiamato file .REC, a partire dal questionario. ENTER legge un questionario già prodotto con un editore di testi in formato testo ordinario e costruisce i nomi delle variabili a partire dal testo del questionario, derivando i simboli e la lunghezza delle variabili dalle informazioni presenti nel questionario stesso. Dopo aver creato un file di dati, ENTER fa apparire il questionario sullo schermo del computer e permette l'inserimento di dati in pochi secondi.

L'editore di testi EPED è provvisto di opzioni particolari per la costruzione di questionari, ma può anche essere usato per la corrispondenza, per fare relazioni e per altri scopi, così come descritto nel manuale elettronico (capitolo 6). Se si preferisce usare un altro editore di testi capace di produrre file di testo semplici (ASCII) oppure usare un file non-documento del Wordstar, è possibile ugualmente produrre questionari utilizzabili con Epi Info. In questo corso sarà tuttavia usato l'editore EPED per costruire i nostri questionari.

IMPOSTARE EPED PER FARE UN QUESTIONARIO

Avviare EPED dal menu EPI6. Premete <F6> per visualizzare il menu SETUP. La prima scelta è tra i tre modi: WW/TXT/QES. Premere la barra spaziatrice una o più volte finché non appare il modo QES, poi premere <Esc>. In questo modo si è stabilita la dimensione delle pagine e altre caratteristiche con i valori adatti ad un questionario.

INDICARE LA POSIZIONE PER L'INSERIMENTO DEI DATI NEL QUESTIONARIO

Per sviluppare un questionario in grado di essere adoperato dal programma ENTER, sono necessarie alcune semplici convenzioni per dire al programma dove creare variabili e campi di immissione dati, e che genere di dati accettare in queste posizioni. ENTER darà dei nomi a questi campi, i quali diventeranno le variabili suscettibili di essere elaborate dal programma ANALYSIS.

In EPED c'è un comando speciale per fare in modo di definire facilmente i campi nel questionario. Per attivarlo, digitare <CTRL - Q Q> (tenere premuto il tasto Ctrl e insieme digitare Q una volta; lasciare quindi il tasto CTRL e digitare ancora una semplice Q). Sullo schermo apparirà un menu contenente vari tipi di campi. Per inserirne uno nel questionario da costruire o aggiornare, muovere l'evidenziatore su una delle scelte e premere il tasto <Invio>. Per alcuni tipi di campi EPED potrà essere richiesta la lun-

ghezza del campo o il numero delle cifre. Appena fornite queste informazioni, il campo sarà inserito nel questionario nella posizione in cui si trova il cursore.

Se per progettare il questionario si sceglie di utilizzare un editore di testi diverso da EPED, si dovrà scrivere da soli l'appropriata serie di caratteri necessari per definire il campo, dal momento che non si avrà a disposizione il comando <CTRL-Q Q>. Di seguito sono mostrati i diversi tipi di campi, i diversi caratteri speciali utili per la loro definizione e la spiegazione del loro significato.

TIPI DI CAMPI IN EPI INFO

—
——
————

Campi di testo, o sottolineati (underline), indicati da caratteri sottolineati in modo continuo. In questo tipo di campo può essere immesso qualsiasi carattere stampabile. I campi vuoti (solo spazi) sono interpretati dal programma ANALYSIS come dati mancanti. La lunghezza della variabile o campo sarà il numero dei caratteri sottolineati usati. La lunghezza massima di un campo di testo è di 80 caratteri.

##.#
ecc.

Campi numerici. Saranno accettati solamente numeri o spazi vuoti. Se non si immette nulla, il risultato sarà in bianco, il che è interpretato da ANALYSIS come un valore mancante e rappresentato con un punto (.). Il numero delle cifre è indicato dal numero dei caratteri "#". Se sono introdotti dei decimali, il campo sarà in formato decimale fisso (fixed decimal), il quale permette esattamente il numero di cifre indicato a destra del punto decimale. Per le somme di denaro da 000.00 a 999.99, per esempio, il formato giusto del campo sarà: ###.##. Il numero da inserire nel campo può essere lungo fino a 14 caratteri, contando anche il punto dei decimali.

<A>
<A >

Campi con lettere maiuscole (upper-case). Questi campi sono simili ai campi sottolineati (underline), solo che i caratteri immessi saranno convertiti in caratteri maiuscoli. La lunghezza del campo è indicata dal numero di caratteri compresi fra le parentesi angolari "<" e ">".

<Y>

Campi "Si/No" ("YES/NO"). Sono accettati soltanto i caratteri Y, N, SPAZIO, e <Invio>. I tasti <Invio> e SPAZIO saranno interpretati da ANALYSIS come dati mancanti. I dati sono convertiti in maiuscolo non appena sono immessi.

<mm/dd/yy>
<mm/dd> m = Month, Mese
<dd/mm/yy> d = Day, Giorno
<dd/mm> y = Year, Anno

Campi per date americane ed europee. Le date saranno controllate al momento dell'immissione per assicurarsi della loro correttezza. Si devono scrivere soltanto i numeri della data, sarà il programma a provvedere a separarli con le barre oblique (giorno e mese devono sempre avere due cifre).

<phonenum> (numero di telefono in formato americano)

Numero di telefono locale, inserito nel formato xxx.xxxx, dove le x stanno per cifre.

<long distance> (interurbane in formato americano)

Un numero di telefono interurbano va inserito in questo modo: (xxx)xxx.xxxx, dove le x stanno per cifre.

<today>

<today/yy>

<today/yyyy>

La data del giorno (today) o data dell'ultimo aggiornamento del campo, inserita automaticamente quando il file è stato salvato l'ultima volta. Se il documento registrato è aggiornato successivamente e salvato di nuovo, il contenuto del campo sarà l'ultima data in cui il documento è stato salvato. Le tre forme permettono di inserire nel campo soltanto il mese e la data; o il mese, la data e l'anno a due cifre, oppure mese, giorno e anno a quattro cifre. La data che appare automaticamente nel campo è la data di sistema del computer, che va inserita ogni volta che lo si accende, a meno che non lo faccia il computer automaticamente. All'interno di Epi Info la data è inserita automaticamente nel campo e non c'è nient'altro da fare che premere il tasto <Invio>.

<IDNUM>

<IDNUM >

Questo campo particolare immette automaticamente dei numeri sequenziali di identificazione. Il primo numero nel file sarà il numero 1 e ogni record successivo avrà un IDNUM (numero di identificazione) maggiore di 1 rispetto al precedente. Questo campo è usato per determinare automaticamente un'identificazione numerica unica, che sarà composta da 5 a 10 cifre a seconda della strutturazione che si desidera.

Dal momento che i campi IDNUM sono riempiti automaticamente, il cursore li salta durante l'inserimento dei dati. Se si desidera fissare il valore del campo <IDNUM> del primo record del file con un numero diverso da 1, basta usare i tasti freccia per entrare nel campo e scrivere il numero desiderato prima di salvare il record.

I segni per le sottolineature, quelli per i numeri (#) e le parentesi angolari, non devono essere usati nel questionario in maniera diversa da quella mostrata sopra, altrimenti il programma penserà che si sta cercando di creare campi scorretti. In particolare non sono permesse le parentesi angolari, né in coppia < > né da sole (ad indicare "maggiore di" e "minore di").

I campi contenenti parole particolari come <today/yy> devono essere esattamente della lunghezza mostrata per contenere correttamente le informazioni sulle date o sui numeri di telefono.

AIUTARE EPI INFO A CREARE NOMI DI VARIABILI SIGNIFICATIVI

Al contrario di molti programmi gestori di basi dati, Epi Info crea automaticamente i nomi delle variabili o dei campi, prelevandoli direttamente dal questionario. Quando il

programma incontra un carattere sottolineato o altri caratteri speciali, cerca sulla stessa linea un testo che possa servire a dare un nome alla variabile. I primi dieci caratteri senza punteggiatura, subito prima del carattere sottolineato, considerati "significanti" da Epi Info, diventano il nome della variabile (Epi Info cerca di evitare le parole inglesi "insignificanti" come "what" "who" "the" "of" "and" "are" (quale, chi, il, di, sono). Il nome della variabile risultante è quello che si dovrà usare per riferirsi a un campo lavorando con il programma ANALYSIS. Per esempio, se la linea è:

Nome _____

il campo sarà chiamato NOME. Se è:

Come ti chiami? _____

il campo sarà chiamato COMETICHIA, dal momento che sono scartati automaticamente gli spazi e il punto interrogativo. Un primo tipo di controllo che si può esercitare sui nomi delle variabili è quello di mettere nomi significanti nei primi dieci caratteri di testo immediatamente precedenti il carattere speciale designante il campo. Ma Epi Info fornisce anche un metodo più preciso per ottenere un determinato nome di variabile: circondando i caratteri desiderati con parentesi graffe { }.

Le parole comprese tra parentesi graffe saranno preferite, come nomi di variabile, rispetto a quelle che sarebbero state generate normalmente. Così, come ti {chiami}, farà introdurre nel campo il nome CHIAMI piuttosto che COMETICHIA.

Si può anche usare più di una coppia di parentesi graffe, a patto che siano sulla stessa linea. Per Esempio: {come} ti {chiami}, dà come nome di campo COMECHIAMI.

Se c'è più di uno spazio vuoto sulla stessa linea, Epi Info comincia la ricerca di un nome di campo utile dopo lo spazio vuoto precedente. Così da:

Nome _____ **Sesso** _____ **Eta** _____

si producono tre campi chiamati NOME, SESSO, ETA (E' consigliabile non usare le lettere accentate). Ora inizia la parte migliore: si supponga che non ci sia del testo precedente il campo sulla stessa linea, come in:

Indirizzo _____

Epi Info chiamerà questi due campi INDIRIZZ e INDIRIZZ01. Il programma rileva anche i duplicati come l'occorrenza di un altro

Indirizzo _____

molte linee più avanti rispetto alle precedenti due. Piuttosto che attribuire un nome di campo doppione, il programma chiamerà questo campo INDIRIZZ02. Il programma ENTER può trattare fino a 99 simili ripetizioni di un singolo nome di campo. Se si desidera numerare i campi lo si può fare in questo modo:

1. Nome _____

ma per prevenire problemi ai programmi che richiedono che un nome di campo inizi con una lettera, Epi Info inserisce una "N", prima dei numeri, se questi sono il primo carattere in un nome di variabile. Si può tuttavia evitare questo scrivendo:

Nome 1. _____

È possibile cambiare i nomi dei campi anche dopo la creazione di un file di dati. Si veda per questo il capitolo dedicato al programma CHECK.

LABORATORIO: EPED

Avviamo Eped : posizioniamo la barra evidenziata su EPED, nel menu PROGRAMMI:
Appare per un attimo il nome del programma e quindi ci troviamo in una pagina vuota, con in alto una riga di comandi e in basso una riga di informazioni sullo stato di lavoro (barra di stato).

Nella prima riga in alto troviamo i tasti funzione :

I tasti funzione aprono delle finestre nella quale sono contenuti dei comandi per attivare il comando occorre posizionarsi sopra con il cursore premere invio, per uscire dalla finestra premere esc.

F1 help = tasto di aiuto, funziona in tutti i programmi di Epi Info

F2 File = per aprire il menù di gestione dei file

(open file this window = aprire il file in questa finestra)

F3 Epiaid = apre un ulteriore programma di Epi Info. Si tratta di un corso di autoistruzione su EPED per utilizzarlo come editore di testi.

F4 = TXT = per gestire i testi possiamo scegliere il tipo di campi di immissione dati in funzione del genere di dati da accettare in queste posizioni.

F4 = QUESTIONS è un menù che contiene vari tipi di campi .

Per i campi carattere è richiesta la lunghezza del campo .

Per i caratteri numerici il n° delle cifre e dei decimali.

F5 = Per accedere alla stampante

F6 = Per settare i parametri dei colori, margini e formato del documento.

F7 = Funzione di ricerca

F8 = Manipola i blocchi di caratteri per copiarli o spostarli.

F10 = Torna al menu principale

Nella **BARRA DI STATO** (l'ultima riga in basso) troviamo alcune informazioni: il numero 1 indica il numero di finestra in cui ci troviamo (questo sta ad indicare che è possibile aprire più finestre cioè più documenti contemporaneamente).

UNTITLED = non intitolato, che sarà sostituito dal nome del file su cui stiamo lavorando o che andremo a creare (il nome del file non può essere più lungo di otto caratteri).

Poi troviamo il numero di byte disponibili nella memoria del Computer (la dimensione del file più grande che può essere caricato)

Ln = numero della riga in cui ci troviamo con il cursore in quel momento.

Cn = numero della colonna in cui ci troviamo con il cursore in quel momento.

(spostandoci con il cursore infatti cambia il numero sia della colonna andando a destra e a sinistra sia della riga in alto e in basso)

INS = inserimento (premendo dalla tastiera il tasto Ins la scritta sul video cambierà in:

OVR = sovrascrive (questo tasto serve per correggere oppure aggiungere dei caratteri)

ESEMPIO: andiamo a creare un questionario che intitoliamo:

QUESTIONARIO DI PARTECIPAZIONE AL CORSO

Partecipante <idnum> Sesso <A>
Età ## Professione _____
U.S.L. ## _____

Un questionario è composto da tante voci, ogni voce è costituita da

1. una domanda
2. uno spazio per registrare la risposta.

Le domande diventeranno i nomi delle nostre "variabili", che useremo al momento dell'elaborazione.

Dobbiamo quindi fare attenzione, quando creiamo un questionario, a dare dei nomi significativi alle variabili, in quanto Epi Info accetta delle variabili non più lunghe di dieci caratteri.

È possibile, nel caso ci trovassimo con una variabile troppo lunga, racchiudere la parte più significativa tra due parentesi graffe; in questo modo il programma leggerà solo le lettere tra le parentesi. Per esempio con {ASS}ISTENZA {DOM}ICILIARE creeremo la variabile ASSDOM.

Abbiamo detto che il questionario è composto da tante voci; ogni voce è una domanda che comporta uno spazio per la risposta, che noi andremo a codificare.

Premendo F4 oppure il tasto <CTRL Q-Q> ci troviamo di fronte a diversi tipi di campi.

1. _____ simbolo sottolineato **campi di testo** alfa numerico (può essere immesso qualsiasi tipo di carattere).
2. # **campi numerici** accetta solo numeri, il numero delle cifre è indicato dal numero dei caratteri, ogni cancelletto è una cifra.
3. <A > solo caratteri **maiuscoli** simile al sottolineato ma solo per lettere maiuscole.
4. **campo logico** <Y> accetta solo Y/N.
5. **Campi per date**, che possono essere **americane** (mese/ giorno/ anno) oppure **europee** (giorno/ mese /anno) , oppure parziali (mese giorno, giorno, giorno/mese).
6. <Phonenum> **Numero di telefono**
7. <long distance> **Prefisso**, da usare solo in caso di prefisso americano, che è molto lungo.
8. <Today> data del giorno.

<Idnum> numeri di identificazione sequenziali (Epi Info immette automaticamente il numero) N° di scheda.

F9 Per salvare il questionario o documento. Viene richiesto il nome, che non deve contenere più di otto caratteri, seguito dall'estensione .QES.

Salviamo il nostro questionario con il nome CORSO.QES.

Possiamo verificare se il computer ha salvato il file: al posto di UNTITLED: in basso a destra ci sarà scritto il nome del file CORSO.QES.

Per ulteriori dettagli, si vedano i capitoli 6 e 7 del Manuale elettronico.

Capitolo 5

COSA BISOGNA SAPERE SULLE BASI DI DATI

LA DEFINIZIONE DI UNA BASE DI DATI E LE SUE COMPONENTI FONDAMENTALI

Una base di dati è fondamentalmente una raccolta di informazioni memorizzate su supporto magnetico. Da una raccolta di informazioni ci aspettiamo, tuttavia, qualcosa di più di un semplice archivio per recuperare documenti. Immaginiamo un grande armadio nel quale riponiamo, giorno dopo giorno, tutti i documenti dell'ufficio: come fare per ritrovare un documento riposto qualche anno fa e accatastato chissà dove fra i mucchi di carta? Nel riporre i documenti dobbiamo evidentemente organizzare le carte, in qualche maniera, in modo tale che ciascuno di essi possa essere recuperato al momento opportuno.

Questo fatto ci introduce alla caratteristica principale di una base di dati che è l'**organizzazione** delle informazioni. Una base di dati, più che un contenitore, deve essere un sistema di smistamento delle informazioni, una struttura capace di accogliere i dati in maniera ordinata e omogenea.

Nei paragrafi che seguono vengono presentati i principali strumenti per organizzare le informazioni in una base di dati. Gli elementi di una base di dati possono essere raggruppati in quattro categorie, a ognuna delle quali sarà dedicato un paragrafo. Le categorie sono:

- 1. Tabelle:** Sono i mattoni della base di dati. In una o più tabelle, composte da righe e colonne, sono registrate le informazioni "brute" dell'intera base di dati.
- 2. Relazioni fra tabelle:** Per far funzionare assieme due o più tabelle è necessario mettere le tabelle in relazione fra loro. Ciò comporta una particolare attenzione nella costruzione delle tabelle.
- 3. Interrogazioni:** Un sistema di tabelle non è ancora una base di dati. E' necessario ancora prevedere una serie di interrogazioni (query), una serie di comandi che eseguono calcoli e filtri sui dati, per estrarre i risultati che interessano.
- 4. Presentazione dei risultati:** Infine, una base di dati, deve prevedere una serie di schemi per inserire i dati e per presentare i risultati. Si deve approntare una serie di maschere per inserire i dati, per mostrare a video dei risultati e per stampare dei tabulati.

LE TABELLE

Le tabelle sono i componenti fondamentali delle basi di dati. Una tabella è una lista organizzata in righe e colonne. In ciascuna riga sono registrate le informazioni su un dato oggetto secondo caratteristiche predefinite, una caratteristica per ogni colonna. Nel linguaggio tecnico, le righe sono chiamate **record** (registrazioni), e possono essere in numero indefinito, mentre le caratteristiche dell'oggetto che si vogliono registrare, le colonne, sono i **campi** della tabella. L'elenco del telefono è un ottimo esempio di tabella:

Cognome	Titolo	Nome	Commento	Indirizzo	Frazione	Telefono
Ancelotti	Dott.	Amedeo	Medico chirurgo	48 vl. Abruzzi	Milano marittima	748392
Di Carlo		Carlo		34 v. Mimose		254784
Guadagni	Rag.	Piero	Studio commerciale	23 c. Mazzini		876657
Rigoni	Ing.	Carlo		45 v. Garibaldi	Savio	875698

I passi da compiere per organizzare una tabella sono due:

1. Individuare l'**oggetto** di cui si vogliono registrare le informazioni (unità di rilevazione)
2. Individuare le **caratteristiche**, gli attributi da rilevare dell'oggetto.

Nel caso dell'elenco del telefono l'oggetto della tabella sono gli abbonati al telefono, mentre le caratteristiche descrittive dell'oggetto sono i sette attributi riportati nella prima riga. Ciò che rende consistente la tabella è il fatto che ogni abbonato registrato è descritto secondo attributi omogenei e fissi. In un secondo momento sarà così possibile compiere operazioni sulla tabella; per esempio raggruppare tutti gli abbonati che hanno il titolo di dottore, oppure ordinare la lista in ordine alfabetico sul campo cognome, o infine selezionare gli abbonati che usano il telefono per lavoro (campo "commento" non nullo).

ORGANIZZAZIONE

Nel sistema EPI INFO il compito dell'organizzazione delle informazioni è ricoperto dal questionario. Ad ogni domanda contenuta nel questionario corrisponde un campo della tabella, mentre ogni questionario compilato costituisce un record della tabella. In effetti quando poniamo una domanda su un oggetto o su un individuo, non facciamo altro che selezionare una caratteristica, un attributo di quell'oggetto o di quel soggetto. Per esempio al questionario che segue corrisponde la tabella subito sotto.

Questionario

Nome _____ Cognome _____

Nato il _____ A _____

Sesso __ Malattia _____

Nome	Cognome	Nato il	A	Sesso	Malattia
Mario	Tizio	24/04/44	Napoli	M	Alfa
Maria	Caio	23/12/52	Ravenna	F	Beta
Carlo	Sempronio	12/10/65	Bologna	M	Gamma

In Epi Info ci sono due programmi distinti per gestire le tabelle. Il programma EPED permette di costruire i questionari, mentre il programma ENTER serve per inserire i record nella tabella. Si rimanda pertanto al manuale di Epi Info per le notizie sull'uso di questi programmi.

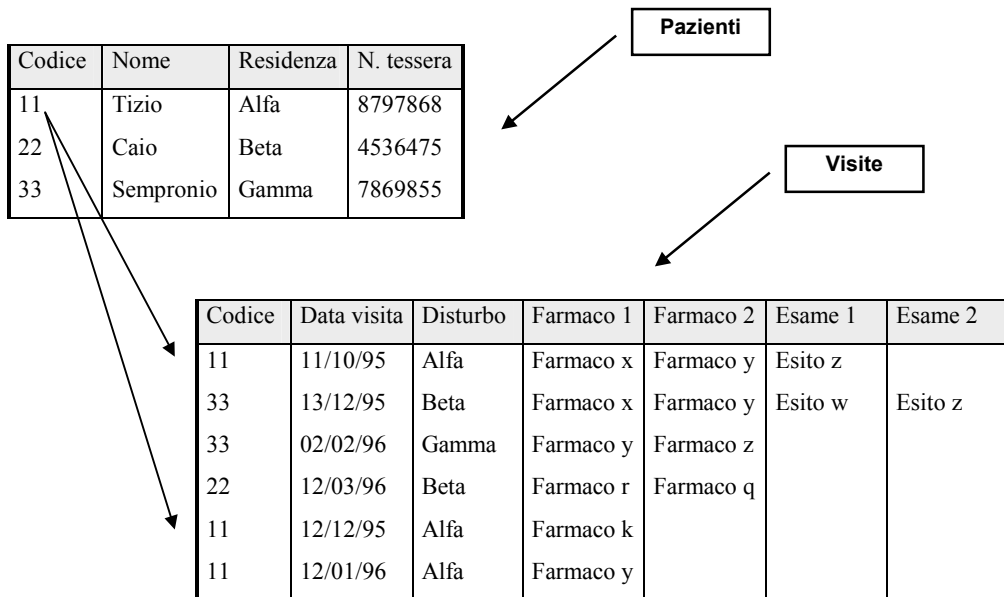
RELAZIONI FRA DUE TABELLE

Una base di dati può essere composta da una sola tabella, con un numero anche grande di campi e record. Di solito, però, si preferisce creare più tabelle in relazione fra loro. Un'altra caratteristica fondamentale delle basi di dati, infatti, è la possibilità di **mettere in relazione** fra loro due o più tabelle.

Un esempio chiarirà meglio il concetto di relazione fra tabelle. Supponiamo di dover creare una base di dati che raccolga le informazioni sulle successive visite di un medico di base ai propri pazienti. Una prima tabella conterrà le notizie su ogni paziente: nome, residenza, età, sesso, numero del tesserino sanitario, ecc. Ma come fare per archiviare le notizie su ogni visita di ogni paziente? Utilizzando un'unica grande tabella, questa apparirebbe, semplificando, come segue:

Nome	Residenza	N. tessera	Visita 1	Visita 2	Visita 3	Visita4	Visita 5
Tizio	Alfa	8797868	Farmaco x	Farmaco y	Farmaco z		
Caio	Beta	4536475	Farmaco x	Farmaco x	Farmaco x	Farmaco q	Farmaco r
Sempronio	Gamma	7869855	Farmaco n	Farmaco p			

Il difetto di questa tabella è che il paziente Caio, alla sesta visita, costringerà il medico di base a ristrutturare la tabella aggiungendo una nuova colonna. Inoltre, se si vogliono registrare più informazioni su ogni visita, bisognerà moltiplicare i campi. Per esempio "Data visita 1", "Disturbo visita 1", "Farmaci visita 1", "Data visita 2", "Disturbo visita 2", e così via.. Il modo per ovviare a queste difficoltà è creare due tabelle, la prima con le notizie generali sui pazienti, la seconda con le notizie di ciascuna visita. Le due tabelle saranno messe poi in relazione.



Per mettere in relazione due tabelle è necessario che entrambe contengano un campo con un codice identificativo univoco in comune. Le due tabelle sopra hanno in comune il campo "Codice". Come si vede, il paziente 33 è stato visitato due volte, il paziente 11 tre volte per lo stesso disturbo, e il paziente 22 una sola volta.

I software che gestiscono le basi di dati funzionano in questo modo: caricano in memoria entrambe le tabelle, leggono la tabella principale (Pazienti) e, ad ogni record, individuano le occorrenze del codice di quel record (in comune con l'altra tabella) nella tabella subordinata (Visite). Così quando si visualizza il record del paziente Sempronio (codice 33) il programma va a cercare tutti i record della tabella visite con il codice 33. In questo modo si possono facilmente creare le cartelle cliniche di ogni paziente creando delle **schede** a video o dei **report** a stampa (si vedano più avanti le definizioni di scheda e report).

RELAZIONI FRA PIÙ TABELLE

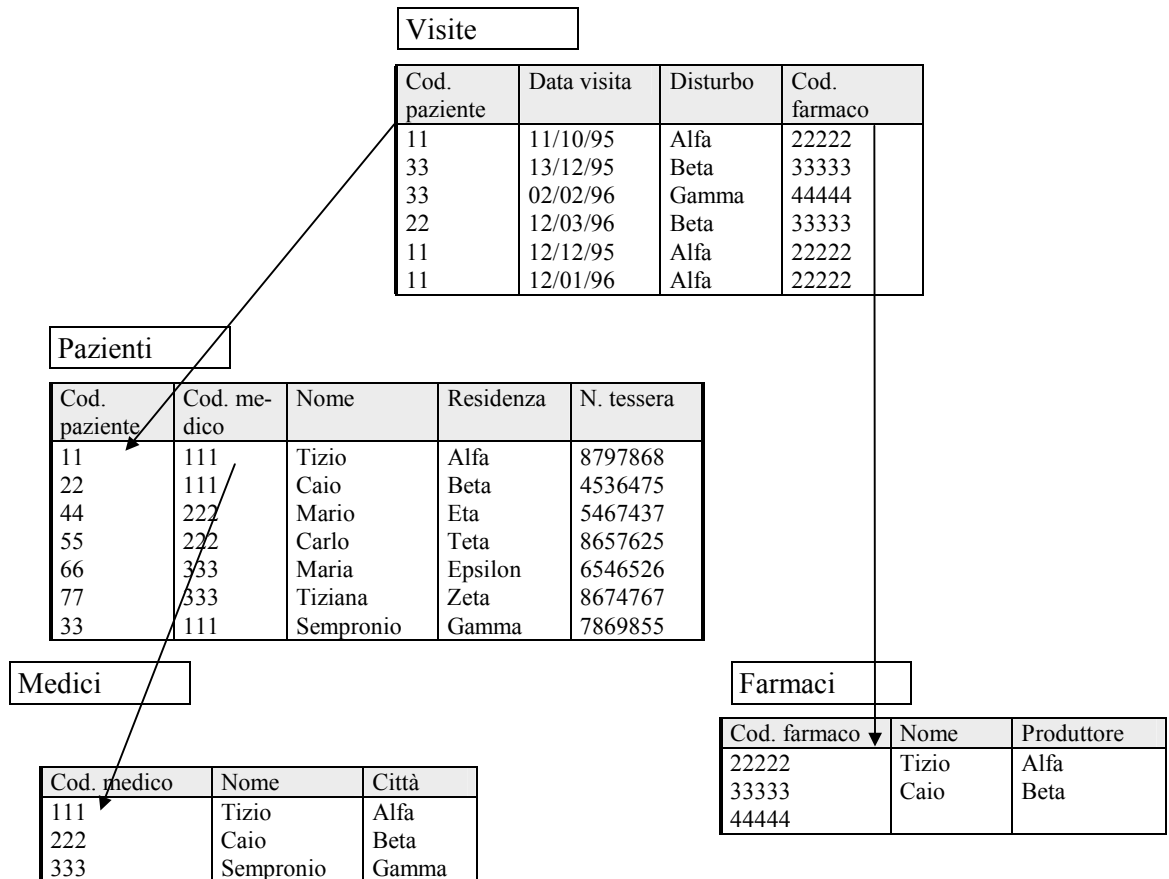
Con la stessa logica con la quale si sono messe in relazione due tabelle è possibile mettere in relazione n tabelle, a seconda della potenza del software impiegato. Il sistema delle relazioni fra tabelle è lo scheletro, l'organizzazione della base di dati. Nella pagina successiva si trova un esempio di relazioni fra quattro tabelle che potrebbe essere utilizzato per controllare la somministrazione dei farmaci da parte dei medici di base.

Seguendo il tracciato delle frecce è evidente il modo in cui è possibile recuperare le informazioni. La tabella principale è quella delle visite. Nel primo record troviamo il codice paziente "11", che rimanda alla tabella pazienti, dove troviamo tutte le informazioni sul paziente. Nella tabella pazienti troviamo che il medico del paziente 11 ha il codice "111", collegato alla tabella dei medici, dove troviamo tutte le informazioni sul medico. Tornando alla tabella visite, troviamo ancora un codice, quello dei farmaci somministrati. Il rimando è alla tabella dei farmaci.

In generale, a partire dalla tabella delle visite troviamo un codice paziente che rimanda alla tabella pazienti, da cui ricaviamo le informazioni sul paziente e il codice del suo medico che rimanda alla tabella dei medici. Oltre alla data della visita e al disturbo, no-

tizie che possono essere articolate a piacere, troviamo il codice del farmaco somministrato, che rimanda alla tabella farmaci.

Come si vede, con quattro tabelle in relazione, composte di pochi campi, è possibile evitare un unico database che richiederebbe centinaia di campi. D'altra parte l'utilità delle tabelle in relazione sta anche nella flessibilità con la quale è possibile interrogare la base di dati per estrapolare informazioni di insieme. Su questo argomento si veda il paragrafo seguente sulle "interrogazioni".



Base di dati con quattro tabelle in relazione

INTERROGAZIONI

Una base di dati non è soltanto una mera raccolta di dati. Abbiamo già visto che la costruzione di tabelle richiede una strutturazione in categorie (le colonne o campi) che costringono a compiere delle scelte. Le tabelle inoltre devono essere messe in relazione fra loro per organizzare meglio la raccolta dei dati. Ma la raccolta ben strutturata dei dati non è un fine, bensì un mezzo per arrivare alla vera finalità di una base di dati: quella di rispondere a quesiti sulla realtà cui si riferisce la base di dati.

La predisposizione di strumenti di interrogazione (in inglese l'interrogazione di basi di dati si chiama appunto *query*, quesito) è la terza delle categorie di componenti di una ba-

se di dati. In Epi Info tutte le funzioni di interrogazione sono svolte dal programma Analysis, il quale dispone anche di sofisticati schemi di calcolo statistico, non presenti nei normali software per basi di dati. Ciò che importa in questo capitolo è capire la logica con la quale operano i sistemi di interrogazione, più che addentrarsi nei comandi specifici di Analysis, descritti con ampiezza nel relativo manuale d'uso.

Possiamo definire una query (d'ora in poi useremo il termine tecnico inglese, ampiamente diffuso) come una ricerca di informazioni eseguita su una o più tabelle collegate. Partendo dall'esempio di base di dati del capitolo precedente, sul controllo di somministrazione di farmaci, potremmo voler estrarre una serie di informazioni. Qual è il farmaco più prescritto, la quantità media dei farmaci prescritti per ogni visita, il consumo di farmaci in rapporto all'età dei pazienti, la spesa farmaceutica media pro capite annua per gruppi di età, la spesa farmaceutica per città o per quartiere, e così via. Tutti questi quesiti sono esempi di query le cui risposte non sono già contenute nelle tabelle costruite, per quanto particolareggiate possano essere: bisogna ricavarle.

Possiamo suddividere le query in tre classi, a seconda del compito che svolgono, tenendo presente che in una singola query possono essere presenti tutti e tre i tipi distinti. Le tre classi sono le seguenti:

1. **Query di selezione:** comprendono le operazioni che servono semplicemente per elencare solo alcuni fra i record che interessano in base al contenuto dei campi.
2. **Query che eseguono calcoli:** comprendono le operazioni che prendono il contenuto di alcuni campi di ogni record e eseguono somme, medie, statistiche ecc.
3. **Query di presentazione:** comprendenti le operazioni per selezionare solo alcuni campi, anche in tabelle diverse, per visualizzare in una sola tabella o scheda, i dati distribuiti in più tabelle

Vediamo ora più da vicino ciascuna di queste tre classi. Si tenga presente che il nostro scopo è quello di **capire la logica** con cui si operano le selezioni, indipendentemente dai comandi particolari che ogni software per basi di dati utilizza per eseguirle.

SELEZIONI MEDIANTE OPERATORI LOGICI

Selezionare solo alcuni fra i record di una tabella è l'operazione più elementare, e di gran lunga la più usuale, quando si interroga una base di dati. Supponiamo di aver raccolto informazioni per migliaia di record. Cercare un nominativo particolare scorrendo tutti i record sarebbe un compito disperato. Il computer si fa carico di questo compito, ma ha bisogno di istruzioni precise sul **cosa** e **dove** cercare. Il "dove" si riferisce solitamente al nome di una tabella, o di più tabelle in relazione, mentre il "cosa" richiede una **descrizione** delle informazioni da cercare.

Una descrizione delle informazioni da cercare non è altro che una serie di attributi legati tra loro da connettivi. Per esempio in: "trova un individuo di sesso X e di età Y, e che sia nato nella città Z", abbiamo una ricerca basata su tre attributi - X, Y e Z - collegati fra loro dalla congiunzione "e", a significare che gli individui da cercare devono possedere tutti e tre gli attributi. Mantenendo gli stessi attributi, ma usando la congiunzione "o", troveremmo un gruppo più ampio di individui, tali che o possiedono l'attributo X, o l'attributo Y o l'attributo Z.

In una base di dati gli attributi non sono altro che i campi di una tabella. Quindi, gli individui che hanno gli attributi X, Y e Z, non sono altro che i record che hanno il valore X nel campo SESSO, il valore Y nel campo ETA' e il valore Z nel campo NATO A. Analogamente, nell'esempio con la congiunzione "o", il risultato della ricerca è quel gruppo di record che hanno almeno uno dei valori X, Y, Z, nei rispettivi campi.

Nel linguaggio delle basi di dati, le congiunzioni che connettono fra loro gli attributi della ricerca sono chiamati **operatori logici**. Abbiamo già visto i due operatori principali, "e" e "o"; ne possiamo aggiungere un terzo: "non". Di solito i nomi degli operatori sono in lingua inglese, anche nei software in versione italiana; così a "e" corrisponde "and", a "o" corrisponde "or" e a "non" corrisponde "not". Ecco di seguito il significato esatto di questi tre operatori:

- AND** Tutti gli attributi connessi devono essere veri. Per esempio RESIDENZA="Ravenna" **AND** SESSO="M", significa che i record da selezionare devono avere **sia** il valore "Ravenna" nel campo RESIDENZA, **sia** il valore "M" nel campo SESSO.
- OR** Almeno uno degli attributi connessi deve essere vero. Nell'esempio sopra, i record da selezionare sono quelli con SESSO = "M" **oppure** quelli con RESIDENZA = "Ravenna", **oppure** quelli con SESSO = "M" e RESIDENZA = "Ravenna" (attenzione, almeno uno non significa "solo uno").
- NOT** L'attributo cui è premesso NOT deve essere falso. Per esempio NOT RESIDENZA="Ravenna" seleziona tutti i record in cui il campo RESIDENZA è diverso da "Ravenna".

Esistono anche altri tipi di operatore, ma variano da un software all'altro, per cui è più utile rivolgersi al relativo manuale d'uso. E' importante invece un accenno alla **concatenazione** di operatori. Possiamo infatti utilizzare più di un tipo di operatore nella stessa ricerca. Per esempio: cerca gli individui di sesso femminile, residenti a Ravenna o a Faenza, che sono coniugati ma non hanno figli. La traduzione nel linguaggio logico è la seguente:

SESSO="M" **AND** (RESIDENZA="Ravenna" **OR** RESIDENZA="Faenza") **AND**
CONIUGATO="Si" **AND** NOT FIGLI="Si"

Come si vede la traduzione rispecchia grosso modo il linguaggio naturale. E' buona regola formulare prima in lingua naturale la ricerca da compiere e poi tradurla sostituendo i nomi dei campi agli attributi della descrizione e gli operatori alle congiunzioni. Le parentesi, infine, seguono le stesse regole che hanno in aritmetica. Bisogna prima compiere le operazioni fra le parentesi più interne e poi via via in quelle più esterne.

CALCOLI

In alternativa o in concomitanza della selezione di alcuni record fra molti è possibile compiere diverse operazioni di calcolo. Per esempio contare quanti sono i record risultanti dalla selezione mostrata sopra, oppure calcolare la percentuale degli individui di sesso maschile sul totale, o ancora sommare o fare la media dei campi numerici (per esempio il numero di figli o le età). Ogni software di basi di dati mette a disposizione un gruppo di operazioni di calcolo da applicare ai record selezionati. In più, Epi Info, offre sofisticati strumenti statistici peculiari alle indagini epidemiologiche.

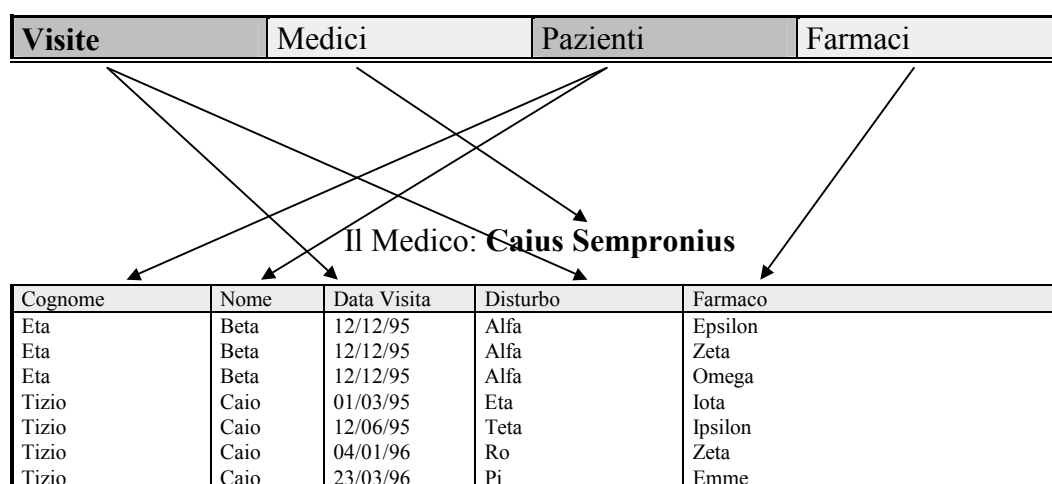
Tutte queste operazioni rientrano nel campo delle interrogazioni e non si può che rimandare alle istruzioni specifiche di Epi Info per maggiori dettagli. Si veda in proposito l'opuscolo sui comandi di Epi Info e, in particolare, i comandi di ANALISYS, il programma di interrogazione di Epi Info.

PRESENTAZIONI

Un'ultima funzione delle interrogazioni è quella di raccogliere i dati in diverse tabelle, selezionare i soli campi che interessano per un determinato scopo, e riunirli in un'unica tabella per **presentare** dei risultati in maniera più leggibile e intuitiva.

Torniamo al sistema di quattro tabelle in relazione fra loro presentato precedentemente. Potremmo voler costruire un'unica tabella, rappresentante la cartella clinica di ogni paziente, in modo che un medico di base la possa utilizzare, durante una visita, per avere un quadro completo della storia del paziente.

Mediante i comandi opportuni, è possibile ottenere la seguente tabella che trae i suoi campi da ciascuna delle quattro tabelle di origine:



Questa tabella unica non è utile in sé, per essere consultata o stampata. Infatti contiene parecchie ripetizioni, per esempio il nome e il cognome del paziente ad ogni sua visita e per ogni farmaco somministrato. E' possibile però **raggruppare** i dati ripetuti, come è stato fatto per il campo MEDICO, che compare una sola volta (il nome del medico, infatti, comparirebbe per tutti i record).

L'utilità della tabella è quella di poter essere usata come base per una **scheda** o un **report** dove è possibile compiere i raggruppamenti del caso. Per questi nuovi elementi di una base di dati, schede e report, si veda il capitolo seguente sulle visualizzazioni.

VISUALIZZAZIONI

L'ultima componente essenziale di una base di dati è quel sistema di oggetti che permettono di visualizzare e di stampare o alcune interrogazioni di tabelle, o delle tabelle inte-

re. Quando una tabella contiene molte colonne e molti record, è difficile, e anche scomodo, scorrere lo schermo a destra e sinistra o stampare tabulati enormi. Lo scopo di schede, report e etichette è quello di sistemare i dati, sullo schermo o su carta, per produrre schermate intuitive per l'inserimento dei dati, moduli cartacei per la presentazione di alcuni risultati di query, o semplicemente per stampare le etichette postali da incollare sulle buste.

SCHEDE

Una scheda è una schermata a video dove vengono sistemati in forma elegante i campi di un solo record alla volta. Con i tasti di cambio pagina è poi possibile scorrere la tabella andando al record precedente o successivo.

In Epi Info le schede sono create da EPED. Quando si scrive un questionario, in realtà si sta costruendo una scheda nella quale ENTER inserirà i valori dei campi della tabella, un record alla volta, al posto dei codici di campo inseriti nel questionario. La schermata di immissione dati, quando si avvia ENTER, è una scheda.

A seconda del software impiegato, le schede possono essere impiegate anche per presentare dei risultati a video, sempre un record alla volta, e non solo per l'inserimento dei dati: per esempio per consultare la cartella clinica di un paziente.

REPORT

I report sono, come le schede, uno strumento per presentare i risultati delle interrogazioni della base di dati. La differenza sta nel fatto che i report sono stampati su carta e richiedono una serie di accorgimenti per presentare i risultati.

Abbiamo visto che non è utile stampare direttamente su carta un'intera tabella, sia per le ripetizioni che contiene e soprattutto per le dimensioni. E' necessaria allora una query preliminare per selezionare i dati da stampare. Una volta fatto questo inizia il lavoro vero e proprio della costruzione del report.

Bisognerà provvedere a inserire un'intestazione, una numerazione delle pagine, le descrizioni dei campi e la disposizione dei risultati sulla pagina. Tutte queste operazioni sono gestite da Analysis e si rimanda al manuale d'uso per le specifiche del caso.

Importa invece menzionare una delle funzioni più specifiche dei report: quella del raggruppamento dei dati. Abbiamo visto in precedenza che per costruire una cartella clinica si è dovuto utilizzare una query per raccogliere i dati da quattro tabelle. Il risultato era una tabella contenente diverse ripetizioni.

Un report risolve il problema delle ripetizioni ed è capace di presentare i dati in forma semplice ed elegante. Ecco di seguito come un report raccoglie i dati dalla query costruita in precedenza per sistamarli in forma di cartella clinica.

Come si vede nel Report di seguito, il nome del paziente viene scritto una sola volta, e anche la data della visita, se è stato somministrato più di un farmaco. Confrontando il report con la query diventa intuitivo il vantaggio dei report in economicità e leggibilità.

Studio Dott. CAIUS SEMPRONIUS - Via delle mimose 14 - 48100 Ravenna

Paziente: Eta Beta

<i>Visite</i>	12/12/95	<i>Disturbo</i>	Alfa	<i>Farmaci</i>	Epsilon Zeta Omega
---------------	----------	-----------------	------	----------------	--------------------------

Paziente: Tizio Caio

<i>Visite</i>	01/03/95 12/06/95 04/01/96 23/03/96	<i>Disturbo</i>	Eta Teta Ro Pi	<i>Farmaci</i>	Iota Ipsilon Zeta Emme
---------------	--	-----------------	-------------------------	----------------	---------------------------------

Pagina 1

ETICHETTE

Le etichette non sono che un tipo di report particolare. In effetti è sufficiente impostare i campi vicini fra loro in modo da formare un tipico indirizzo da busta. Di solito i dati vengono raccolti da una sola tabella, e la costruzione di etichette risulta estremamente semplice. Epi Info ha dei comandi specifici per la creazione automatica di etichette: si veda l'esempio contenuto nel menu Esempi di Epi Info alla voce Etichette postali.

Capitolo 6

INSERIRE I DATI USANDO IL PROGRAMMA ENTER

INTRODUZIONE

Il programma ENTER è il gestore delle basi di dati di Epi Info e la sua funzione fondamentale è quella di permettere l'inserimento dei dati dei questionari compilati.

Una volta creato un questionario, ENTER trasforma il testo del file .QES in una maschera video, sostituendo i codici dei campi in spazi gialli dove l'utente dovrà immettere i dati. L'insieme degli spazi della maschera rappresentano un **record**, una riga, che ENTER andrà a immettere in forma tabellare nel file della base di dati.

AVVIARE ENTER

Si sposti il cursore su PROGRAMMI del menu principale di Epi Info e poi su ENTER, e premere <Invio>. Sullo schermo apparirà la domanda circa il file di dati da aprire e mostrerà il seguente menu:

- | | |
|---|---|
| 1. Enter or Edit data | (Inserire o modificare i dati) |
| 2. Create a new data file from .QES file | (Creare un nuovo file dati da un file .QES) |
| 3. Revise structure of data file using revised .QES | (Ristrutturare un file dati con un nuovo file .QES) |
| 4. Reenter and verify records in existing data file | (Reimmettere i dati in un file esistente per verificarli) |
| 5. Rebuild index files(s) specified in .CHK file | (Reindicizza i file indice specificati nel file .CHK) |

Le opzioni 4 e 5 riguardano funzioni avanzate introdotte con la versione 6 di Epi Info, per le quali si rimanda al manuale elettronico in linea. La funzione 4 serve per un secondo operatore che reimmetta gli stessi dati immessi da un altro operatore, per controllare che i dati siano stati immessi correttamente. L'opzione 5 riguarda la manutenzione dei file indice, da utilizzare quando compare un errore di indicizzazione. I file indice sono file contenenti la copia dell'intera colonna di un campo del file dati, ma con un ordine particolare (per es. l'ordine alfabetico dei cognomi dei pazienti). Servono essenzialmente per velocizzare le ricerche di record.

Ciò che qui interessa sono invece le funzioni 1, 2 e 3.

INSERIRE O MODIFICARE I DATI

I file di dati di Epi Info hanno i nomi con estensione .REC, poiché contengono Record di dati. Per inserire nuovi dati o modificare quelli già inseriti in un file (.REC) esistente, digitare il nome del file dati, o premete <F9> per visualizzare quelli disponibili. Scegliete

un file dal menu spostando il cursore con i tasti freccia e poi premere <Invio> per selezionare il file. Digitare "1" per la scelta di menu e premere di nuovo <Invio>. Il file verrà visualizzato sullo schermo e si potrà cominciare subito a lavorare.

CREARE UN NUOVO FILE DI DATI (.REC)

Il file di dati sarà costruito da ENTER a partire dal file .QES del questionario. Nella casella dove viene richiesto il nome del file di dati, scrivere TEST come nome del file che si desidera creare, dato che non esiste ancora un file con questo nome; digitare 2 nella casellina di selezione del menu (Crea nuovo file dati) e premere <Invio>. Non essendoci nessun file con il nome introdotto, ENTER chiede il nome del file del questionario per poter creare un file di dati. Di solito conviene usare lo stesso nome per il file di dati e per il file del questionario, anche se non è affatto necessario. Inoltre non c'è nessun bisogno di scrivere le estensioni .REC e .QES, perché sono date per scontate dal programma.

Nota: Quando il cursore è nello spazio vuoto "Data file" oppure in quello "Questionnaire file", si può premere <F9> per visualizzare tutti i file di dati o di questionari disponibili. Per vedere le disponibilità di un'altra directory o di un altro disco, digitare la lettera e/o la directory del drive di disco e premere <F9>. Per scegliere un file da una lista muovere il cursore con i tasti freccia, portarsi sul file prescelto e premere <Invio>.

Una volta specificati entrambi i file, ENTER impiega pochi secondi a leggere il questionario, compilare un file di dati e mostrare il questionario sullo schermo. A questo punto si è pronti per l'inserimento dei dati.

L'INSERIMENTO DEI DATI

Immettere i dati significa riempire gli spazi vuoti dello schermo con risposte appropriate: il cursore salterà automaticamente da uno spazio vuoto all'altro. La riga dei messaggi, in fondo allo schermo, mostra i comandi disponibili e, nell'angolo a destra, il numero del record corrente. Appena sopra la riga dei messaggi è riportato il tipo di dati che saranno accettati nel campo sul quale è attualmente il cursore. Queste limitazioni possono essere cambiate usando il programma CHECK che si descriverà in seguito; per ora i soli tipi di dati ammessi sono soltanto quelli numerici, le date, un testo generico e gli altri tipi di campi visti precedentemente.

SOMMARIO DEI TASTI FUNZIONE

I tasti funzione sono elencati in fondo allo schermo. Quelli usati per cercare i record si attivano con <Ctrl-F>.

- <F1> Help. Visualizza un file di guida creato per il questionario corrente, se tale file è stato specificato nel file .CHK corrente. In questo caso, il questionario dovrebbe contenere un richiamo Premere <F1> per la guida, in modo che l'utente venga a conoscenza dell'esistenza di questa risorsa. Premendo il tasto <Maiusc-F1> si visualizzerà la guida specifica al campo in cui si trova il cursore.

- <F2> **Rec #.** Trova un particolare numero di record. Si digiti un numero e il record con quel numero apparirà sullo schermo.
- <F3> **1st** (primo). Cerca un record in base al tipo di informazioni contenuto in tutti i suoi campi non vuoti, eccetto i campi <IDNUM> e <Today>. La ricerca parte dall'inizio del file e va avanti fino alla fine. Si può ricercare un record in base a informazioni parziali usando l'asterisco o carattere "jolly" ("SMI*" è un'informazione parziale di "SMITH"); il campo non deve essere però numerico.
- <F4> **Next** (prossimo). Lo stesso che <F3>, eccetto che la ricerca parte dalla posizione corrente nel file e va in avanti. Trova la successiva occorrenza di record rispondente a un determinato requisito
- <Maiusc-F4> **Prev** (precedente). Lo stesso che <F3>, eccetto che la ricerca parte dalla posizione corrente nel file e va indietro. Trova l'occorrenza precedente del record rispondente al requisito ricercato.
- <F5> **Print** (stampa). Permette di stampare il questionario o i dati contenuti nei record specificati. In fondo allo schermo appariranno dei messaggi che guideranno lungo tutto il processo di stampa.
- <F6> **Delete** (cancell). Considera il record corrente come cancellato e sarà pertanto ignorato dal programma ANALYSIS. Se il record è già stato cancellato mediante questo tasto, premendo di nuovo <F6> il record sarà ripristinato. I record cancellati sono indicati da un asterisco nell'angolo in basso a destra dello schermo, vicino al numero di record.
- <F7> **Rec<**. Riporta sul record precedente a quello sullo schermo.
- <F8> **Rec>**. Porta al record successivo, se esiste.
- <F9> **Choices** (scelte). Mostra, se esistono, i codici immessi nel campo corrente per la codificazione automatica. La scelta può essere fatta dal menu che appare, muovendo il cursore con le frecce e premendo <Invio>. Digitando una lettera dell'alfabeto il cursore si porterà sulla prima scelta che comincia con quella lettera. Le scelte che si compiono digitando una lettera sulla tastiera possono essere compiute anche se il menu è stato chiuso premendo <Esc>.
- <F10> **Done**. Esce dal programma ENTER e torna al menu Epi o al DOS.
- <Ctrl-N> **New (nuovo)**. Libera tutti campi nel record corrente e posiziona il cursore nel punto in cui dovrà essere aggiunto un nuovo record al file di dati. Usare questo comando quando si è finito di scrivere sui record e si desidera inserirne di nuovi, oppure per uscire dal modo ricerca (search mode <Ctrl-F>).
- <Ctrl-F> **Find** (cerca). Prepara lo schermo per l'inserimento di criteri per la ricerca di un record.
- <Ctrl-S> **SOUNDEX**. Prepara lo schermo per l'inserimento di criteri per una ricerca di tipo SOUNDEX.

Altri tasti con funzioni particolari sono descritti nelle sezioni seguenti. Quando si inseriscono i dati si tenga presente che:

- I campi accettano soltanto il tipo di dati mostrato sopra la riga dei messaggi.
- Premere <Invio> senza inserire dati significa che il dato è mancante.
- Quando un campo è pieno il cursore si porta automaticamente a quello successivo, a meno che questa opzione non sia stata disattivata (si veda l'inizio del Capitolo 35 del Manuale elettronico di Epi Info).
- Gli errori di immissione dati sono segnalati da un "beep" e possono essere corretti immediatamente.
- Ogni spazio vuoto (o casella) è chiamato "campo". Ogni copia del questionario, completo dei dati, è chiamato "record". I record sono archiviati nei file con estensione .REC, chiamati anche file di dati o file .REC.

Dopo aver inserito i dati nell'ultimo campo di un questionario, in fondo allo schermo appare la domanda: "Write data to disk? (Y/N)" (Scrivo i dati sul disco? Y=sì N=no). Se si risponde "Y" (sì) i dati vengono registrati su un record e ne appare subito un altro da riempire (si noti il cambiamento del numero di record in basso a destra). Se la risposta è "N" (no), il cursore torna al primo campo del questionario dando la possibilità di modificarlo.

Per uscire dal programma, in ogni momento, si può premere il tasto <F10>. Se non si è salvato il record corrente, ENTER chiede se si desidera salvarlo e torna al menu EPI o al DOS.

MUOVERSI FRA UN RECORD E L'ALTRO

Il numero del record corrente è sempre mostrato in basso a destra sullo schermo. Premendo <Ctrl-F> (che sta per "find", "trova") appariranno sullo schermo tutti i comandi che servono alla ricerca e per spostarsi da un record all'altro. Per spostarsi al record precedente si preme il tasto <F7>, mentre con <F8> ci si porta al record seguente, se c'è.

Se, dopo aver "sfogliato" (browsing) il file con <F7> e <F8>, si desidera scrivere su un altro record, tenere premuto il tasto <Ctrl> e premere contemporaneamente "N" (sta per "New", "Nuovo"): lo schermo sarà sgombrato e apparirà il successivo nuovo record disponibile. A volte è necessario premere prima il tasto <Esc> quando si è alla fine del file o se appare un messaggio sullo schermo.

CERCARE RECORD RISPONDENTI AD ALCUNI REQUISITI (MATCHING CRITERIA)

ENTER permette di trovare, in un file di dati, tutti i record che rispondono a determinati requisiti scelti dall'utente. Per compiere queste ricerche ci si assicuri, se necessario, che il file in questione sia stato salvato. Dare il comando <Ctrl-F> (sullo schermo appariranno degli spazi vuoti nei quali scrivere i requisiti in base ai quali cercare i record) e premere <F3>, che sta per "find 1st", "trova il primo" record del file rispondente al requisito introdotto. Per esempio se si scrive Smith nello spazio riservato

al cognome e si preme <F3>, apparirà sullo schermo il record del primo degli Smith presenti nel file, se esistono. Se si vuole andare al successivo "Smith" del file, basta premere <F4> e la ricerca continua a partire dalla posizione del record corrente. Per tornare indietro si preme <MAIUSC-F4> e la ricerca riparte all'indietro dal record corrente.

ENTER procede allo stesso modo anche se si immettono più requisiti, scelti combinando in un modo qualsiasi i campi. Per esempio si potrebbe scegliere "Rossi" come cognome, "Mario" come nome e "47" come età: la ricerca riguarderebbe allora i Mario Rossi di 47 anni.

A volte è utile fare ricerche introducendo nomi parziali, per esempio scrivere "Pedrett" se non si è sicuri che l'individuo cercato si chiami "Pedretti" o "Pedretto". Per farlo si introduce il carattere "jolly" al posto dei caratteri omissi: nel nostro caso si scriverà allora "pedrett*". I caratteri jolly possono essere usati solo in campi non numerici.

Se si ha soltanto qualche centinaio di record la ricerca sarà ragionevolmente veloce. Epi Info contiene però un dispositivo di indicizzazione dei file (indexing facility) che permette di trovare quasi istantaneamente un record fra migliaia di altri. Il modo di creare file indicizzati sarà trattato nel corso avanzato. In ogni caso, dal punto di vista dell'utente, non c'è alcuna differenza nei comandi di ricerca: l'indicizzazione permette semplicemente una ricerca veloce per file molto grandi.

La funzione SOUNDEX è una codificazione fonetica delle somiglianze dei suoni della lingua. Il risultato di una ricerca in base alla funzione SOUNDEX (vedere il cap. 35 del manuale elettronico) comprende, dunque, non solo i record che letteralmente corrispondono a un criterio richiesto, ma anche quelli che *assomigliano* ai quei criteri. Tali ricerche sono molto utili, per esempio, per controllare l'esistenza di un dato cognome in basi di dati di diversa provenienza. Infatti il cognome di uno stesso individuo, potrebbe essere stato scritto in maniera diversa da due diversi operatori. Una ricerca SOUNDEX si esegue premendo <Ctrl-S> piuttosto di <Ctrl-F> per dare il via al processo di ricerca.

MODIFICARE I RECORD GIÀ INSERITI

Per modificare un record già inserito precedentemente, per prima cosa lo si cerca e lo si porta sullo schermo come descritto precedentemente. Quindi, con l'aiuto dei tasti freccia e dei tasti <Canc> e <Ins>, si possono cambiare i dati contenuti nei campi del record immettendo nuovi dati. Si ricordi di premere <Invio> quando si esce da un campo modificato. Una volta compiute le modifiche necessarie bisogna andare all'ultimo campo del record e premere <Invio>, oppure premere semplicemente il tasto <Fine>. Sullo schermo, in basso, appare la domanda "Write data to disk? (Y/N)" ("Scrivo i dati sul disco?"). Se si risponde "Y" (Si) il record corrente rimpiazzerà il vecchio, mentre se si risponde "N" (No) il cursore torna al primo campo del record. Se ci si sposta su altri record con <F7> e <F8> senza salvare le modifiche apportate, queste saranno scartate e il record rimarrà nella sua forma originaria.

CANCELLARE E RECUPERARE I RECORD CANCELLATI

Premendo il tasto <F6> il record corrente è considerato come cancellato. Vicino al numero di record, in basso a destra sullo schermo, apparirà un asterisco (*). Il contenuto del record è ancora visibile nel programma ENTER, ma le tabulazioni compiute dal programma ANALYSIS salteranno questi record a meno che non si istruisca ANALYSIS a considerarli diversamente.

Una volta che il record è stato cancellato, può ancora essere recuperato ("undeleted") premendo di nuovo il tasto <F6>. Questo tasto si alterna dunque nella funzione di cancellazione e ripristino.

RISTRUTTURARE UN FILE DI DATI

È possibile modificare un questionario (file .QES), usando EPED, e creare in seguito un nuovo file di dati in base alle modifiche apportate. Capita spesso di doverlo fare quando si è già introdotto un certo numero di record e ci si accorge di aver tralasciato un attributo importante.

Modificare la struttura di un file di dati deve essere accuratamente distinto dalla semplice modifica dei dati come descritta sopra. Per ristrutturare un questionario bisogna tornare in EPED e modificare il file .QES contenente il questionario originale. Tornati poi in ENTER, si inserisce il nome del file .REC sotto la scritta "Data file:" e si preme <Invio>. Infine si digita "3" come scelta del menu di ENTER e si inserisce il nome del nuovo questionario in base al quale ristrutturare la base di dati.

ENTER costruirà automaticamente il file .REC con la nuova struttura, inserirà nel nuovo file i vecchi dati e cambierà in .OLD l'estensione del vecchio file .REC. L'intero processo è descritto sullo schermo nel corso della sua esecuzione. Questo processo è eseguito correttamente solo se non è stato cambiato il nome dei campi del file .REC. Se si è usato il programma CHECK per cambiare il nome dei campi nel file .REC, ci sarà bisogno di modificare anche il questionario per inserire gli stessi nomi di campo, altrimenti i nomi precedenti dei campi saranno considerati come nuovi campi.

STAMPARE IL QUESTIONARIO E I DATI

Per la stampa di una copia del questionario o di una lista di record, premere <F5>: ENTER chiederà se si vuole stampare una copia del questionario. Se si, digitare "Y", altrimenti scrivere "N". Il questionario sarà stampato con una sottolineatura negli spazi vuoti, pronto per la scrittura a penna.

Per la stampa di una lista di record, viene richiesto da quale numero di record a quale numero di record si vuole la stampa. Ci si assicuri che la stampante sia pronta e si inseriscano i numeri richiesti. Se si scrive zero per entrambi i numeri di inizio e di fine, verrà stampato solo il questionario. Alla domanda: "Single Sheets (Y/N)" ("Fogli singoli S/N"), si risponda "Y" se si vuole una pausa di stampa dopo ogni pagina per cambiare il foglio. Dopo una piccola pausa, ENTER manda i dati alla stampante e i record sono riassunti in forma compressa come segue:


```

RECORD NUMBER 60
ETÀ      :53   SESSO      :F   ORAPASTO  :1930  MALATO    :S
DATAINSORG : 25/02   ORAINSORG :2330  SALAME    :S
SPINACI   :S   PURÈ      :S   CAVOLI    :S   MARMELLATA :S
PANE      :N   PANENERO  :S   LATTE     :N   CAFFÈ     :S
ACQUA     :S   DOLCI    :S   VANIGLIA  :S   CIOCCOLATO :S
MACEDONIA :N

```

LABORATORIO: ENTER

Una volta creato il nostro questionario CORSO.QES, bisogna generare **un file di base di dati** in modo da potere memorizzare i dati in un file .

Avviamo ENTER dalla maschera principale, scriviamo CORSO nella prima riga, e selezioniamo l'opzione 2 (create new .REC file). Nella riga in basso inseriamo il nome del questionario. Con <F9> visualizziamo l'elenco dei questionari e selezioniamo CORSO.QES.

Una volta specificato il nome del file, Enter mostrerà sullo schermo una maschera simile al questionario creato, a questo punto è possibile introdurre i dati.

Inserire i dati in CORSO.REC

Compilare il questionario CORSO.REC consultando il sommario dei tasti funzione. Tenete presente che ogni spazio vuoto è un *campo*, ogni esemplare di questionario completo dei dati sarà chiamato *record*. I record sono archiviati nei file con estensione .REC, chiamati anche file di dati o file rec. o *base di dati*.

Una volta terminato l'inserimento dei dati in fondo allo schermo apparirà la domanda: "scrivo i dati sul disco?" Y/N rispondendo Yes i dati vengono registrati su disco magnetico e appare subito un altro questionario da riempire. Rispondendo No il cursore torna al primo campo del questionario, dando così la possibilità di modificarlo.

Ristrutturare CORSO.REC modificando il questionario

Usciamo da Enter con F10 e torniamo in Eped. Chiamiamo con il tasto F2 il file di nome "CORSO.QES" e aggiungiamo la **data** del corso. A questo punto il questionario è diverso da quello precedente, deve essere quindi ristrutturato il file archivio. Tornando in Enter scriviamo il nome del file .REC, digitiamo la scelta 3 (che come indicato sul video serve per revisionare il File REC), comparirà quindi la richiesta di scrivere il nome del file del questionario che è stato modificato; premendo invio Enter costruirà automaticamente il file .REC con le modifiche apportate e trasferirà nel nuovo file i vecchi dati, quindi cambierà in .OLD l'estensione del vecchio file .REC .

ESEMPIO di navigazione

Entrare nel programma ENTER. Richiamare il file ORIENTE.REC, presente nel quarto dischetto insieme a Epi Info 6, riguardante un'indagine su una tossinfezione alimentare condotta a Ravenna nel 1991.

Per visualizzare quante persone sono state ricoverate, attiviamo CTRL-F, ci spostiamo sulla variabile RICOVERO, digitiamo "Y" e premiamo il tasto <F3> (per il primo record trovato) e con F4 cerchiamo i successivi:

Quanti sono stati ricoverati? = n.4 (record 6 - 69 - 70 - 71)

Vogliamo ora vedere quante persone di 3 anni sono state al Pronto Soccorso:

Avviamo il comando di ricerca CTRL F - ci spostiamo con il cursore sui campi da ricercare (impostiamo età= 3 e PS= Y). Premendo <F3> troviamo il primo record di un bimbo di tre anni che è stato al PS. <F4> per cercare i successivi, <Maiusc-F4> per i precedenti.

Quanti sono i bambini di tre anni che sono stati al pronto soccorso? = n.2
(record 7 - 70)

Per ulteriori informazioni, si veda il capitolo 8 del Manuale elettronico.

Capitolo 7

IL PROGRAMMA CHECK: CONTROLLO DEGLI ERRORI E DEI CODICI

Spesso, nel corso dell'immissione dei dati, è utile che sia il computer stesso a controllare i possibili errori, a decodificare automaticamente i dati in codice o a saltare alcune parti di questionario rispondenti a determinati requisiti: il programma CHECK serve per istruire ENTER a eseguire queste operazioni automaticamente. Usando CHECK per codificare e porre delle restrizioni ai dati da immettere, si possono evitare gli errori più comuni, ma anche rendere più automatica e agevole l'inserimento stesso. Il che è particolarmente utile se sono molte le persone addette a questo compito, e specialmente se hanno una esperienza minima di computer.

CHECK apre un file (con estensione .CHK) che istruisce ENTER a restringere il tipo di dati introducibili in determinati campi. Quando ENTER è avviato, cerca automaticamente un file con lo stesso nome del file .REC caricato, ma con estensione CHK. Usare CHECK è facoltativo: il programma ENTER funziona altrettanto bene anche se si decide di non fare un file .CHK.

AVVIARE CHECK

Prima di avviare CHECK deve già esistere un file .REC corrispondente al questionario che si vuole perfezionare. Se necessario, prima di avviare CHECK, si usi ENTER per costruire un file .REC a partire da un file .QES. Per costruire un file .REC vuoto basta uscire immediatamente da ENTER (con <F10>) senza immettere alcun dato. A questo punto si può avviare CHECK dal menu principale, introdurre il nome del file .REC nello spazio "data file:", e rispondere "Y" (Sì) nello spazio a fianco della scritta: Ready? (Pronto?).

CHECK presenta il questionario sullo schermo e, in basso, le seguenti indicazioni dei tasti funzione:

F1/F2-Min/Max	F5-Collega campi	Legal : F6-Aggiunge	Shift-F6-Mostra	Ctrl-F6 Cancella
F3-Ripete	F9-Modifica campo	Salto : F7-Aggiunge	Shift-F7-Mostra	Ctrl-F7 Cancella
F4-Obbligatorio	F10-Esce	Codice: F8-Aggiunge	Shift-F8-Mostra	Ctrl-F8 Cancella

Appena sopra la serie di tasti funzione, è mostrato il tipo di codici attivi nel campo sul quale si trova il cursore.

I tasti funzione agiscono sul campo su cui è posizionato il cursore. Premendo una volta un tasto funzione si pone una condizione per quel campo e in alcuni casi, premendolo una seconda volta, la condizione viene tolta. Così, se si desidera fare di CITTÀ un campo "Repeat", come descritto più avanti, si deve posizionare il cursore su CITTÀ e premere <F3> (Ripete). Per rimuovere la condizione si preme di nuovo <F3>. La condizione posta in atto è mostrata in basso, sopra l'indicazione dei tasti funzione. In uno stesso campo si possono porre molte condizioni diverse contemporaneamente.

MINIMI E MASSIMI (MIN/MAX) <F1> <F2>

Quando sono specificati un minimo e un massimo, il valore immesso in un campo deve essere maggiore o uguale a un limite inferiore (lower bound o lower range) e minore o uguale a un limite superiore (upper bound o upper range). Ogni tipo di campo può avere dei limiti diversi, anche se il significato è sensibilmente diverso per i campi non numerici. In quest'ultimo caso il nome in formato testo deve stare entro due limiti quando è ordinato in ordine alfabetico.

Per fissare un limite inferiore o uno superiore, scrivere il valore nel campo desiderato e premere <F1> o <F2>. Usando il tasto <F6> si possono aggiungere condizioni per considerare "Legali" anche altri valori oltre quelli compresi fra i limiti. Per esempio, possono essere ammessi i valori fra 1 e 3 e in più il numero 9.

Esempi:

Limite inferiore 1, limite superiore 3: i valori ammessi sono 1, 2 e 3.

Limite inferiore Savini, limite superiore Tognini: "Savini", "Stefanini", "Suzzi", "Tostoni" e "Tognini", sono ammessi, ma "Tognoni" e "Tosi" sono rifiutati.

VALORI PERMESSI (LEGAL) <F6>

I valori permessi sono quelli che saranno accettati come dati nei campi. Quando è specificata una lista di valori legali, un dato in entrata deve corrispondere ad uno di essi, altrimenti sarà rifiutato. Premere <Invio> su un campo per dire che il valore è mancante, sarà accettato come valore "Permesso" (Legal) a meno che nel campo non sia presente la condizione "Obbligatorio" (Must Enter) che sarà esposta più avanti.

Per considerare come permesso un valore, scriverlo nel campo appropriato e premere <F6>. Il valore sarà aggiunto alla lista dei valori permessi. Per eliminare un valore permesso esistente, scrivere il valore sul campo e premere <Ctrl-F6> (cioè tenere premuto il tasto <Ctrl> e premere <F6>). Se i valori sono in numero maggiore di quelli che entrano in basso nello schermo, premendo <Maiusc-F6> (tasto per le maiuscole + F6) sarà mostrata l'intera lista dei valori legali.

INSERIMENTO OBBLIGATORIO (MUST ENTER) <F4>

Quando è specificata l'obbligatorietà (Must Enter), non sono ammessi valori mancanti (campi lasciati vuoti). Per specificare un obbligo di inserimento per un campo, basta posizionarsi sul campo e premere <F4>. Premendo <F4> di nuovo si elimina la condizione di obbligatorietà.

RIPETIZIONE (REPEAT) <F3>

La presenza della condizione "Ripetizione", fa sì che il campo assuma automaticamente il valore del campo corrispondente nel record che lo precede. In questo modo basta introdurre una sola volta un dato nel primo campo e questo sarà ripetuto in tutti i record successivi in quello stesso campo. Ciò è utile quando una serie di questionari

contiene, per esempio, lo stesso nome di città o di regione, o altri dati che non variano durante l'inserimento. Per designare un campo "Ripetizione" posizionare il cursore sul campo e premere <F3>. Premendo di nuovo <F3> si elimina la condizione.

CONDIZIONI PER SALTARE ALCUNI CAMPI (SKIP PATTERNS) <F7>

I questionari contengono frequentemente delle sezioni che possono essere saltate se una risposta a una domanda precedente le rende superflue. Per esempio, se non è insorta nessuna malattia, è inutile andare alle domande sui sintomi e le si potrebbe saltare per andare direttamente al successivo campo utile. È anche utile, a volte, far "saltare" (jump) il cursore su un gruppo specifico di domande se si incontra una certa risposta in un certo campo. Se come risposta per MALATTIA c'è "Rabbia", il cursore potrebbe saltare a una serie di domande specifiche per questa malattia, e tornare alla domanda che segue MALATTIA appena finito.

Per introdurre le condizioni per il "salto" (jump), bisogna immettere un valore nel campo appropriato, per esempio "N" (No) o "Rabbia" nel campo MALATTIA, e premere <F7>. Verrà richiesto di indicare il campo verso cui fare il "salto" se si incontra la risposta "N" o la risposta "Rabbia" in MALATTIA: posizionare il cursore in quel campo e premere <F7> di nuovo. Così, per rimanere con gli esempi dati, se il programma ENTER incontra la risposta "N" nel campo MALATTIA, scavalcherà tutte le domande concernenti il tipo di malattia; oppure, se in MALATTIA è rilevata la risposta "Rabbia", ENTER salterà alle domande specifiche sulla rabbia. Naturalmente se le risposte incontrate sono "Y" (si) o una malattia diversa dalla rabbia il cursore si comporterà normalmente passando al campo successivo.

Per uno stesso campo, se si vuole, si possono specificare molti valori e con condizioni differenti. Per far compiere un salto indipendentemente dal contenuto del campo, premere <F7> quando il campo è ancora vuoto. Questi salti incondizionati sono utili per cambiare l'ordine di introduzione dei dati del questionario.

Per saltare a un certo campo se la risposta del campo è per esempio "1" e a un altro campo se la risposta è diversa da "1", si procede in questo modo: si crea un salto senza condizioni quando il campo è ancora vuoto, poi si indica un altro campo a cui saltare se è presente il valore "1". Naturalmente, nei campi in cui non è stato specificato nessun valore di "salto", il cursore si comporta normalmente andando al campo successivo.

Per poter vedere tutte le condizioni di salto presenti in un campo, si preme <Maiusc-F7>; per rimuoverle si usa <Ctrl-F7>.

CAMPI COLLEGATI E CODIFICA AUTOMATICA <F5> E <F8>

Se in un campo è presente una codifica automatica, il dato in entrata sarà confrontato con una tabella consistente in coppie di codici, parole o frasi, e convertita nel valore dell'altro membro della coppia. Il risultato sarà riportato nel campo stesso o, se è stato specificato un altro campo con il comando di collegamento (Link command), nel campo collegato.

Si supponga di voler collegare due campi, per esempio MALATTIA e CODMALATI, e di voler introdurre dei codici in MALATTIA. Quando si scrive "Epatite B" nel campo MALATTIA, il codice corrispondente, poniamo "10", sarà trovato e riportato automati-

camente nel campo chiamato CODMALATI. E se si digita "10" nel campo CODMALATTIA, automaticamente sarà riprodotto "Epatite B" nel campo MALATTIA.

A meno che non si desideri che i dati immessi siano decodificati e reimmessi nel campo stesso, è necessario collegare (Link) due campi prima di immettere i codici. Per compiere questa operazione di collegamento bisogna portarsi sul primo dei due campi e premere <F5>, quindi spostare il cursore sullo schermo fino al secondo campo e premere di nuovo <F5>: i due campi sono ora collegati e si può procedere all'immissione dei codici. Si supponga di avere di nuovo i campi MALATTIA e CODMALATI e che questi siano stati collegati nel modo già descritto. Quando si immette un valore in MALATTIA, per esempio "Epatite A", e si preme il tasto <F8>, il cursore "salta" a CODMALATI, il campo collegato. Si immetta allora il codice 11, per esempio, e si preme di nuovo <F8>. Il cursore salterà di nuovo a MALATTIA e si sarà pronti per immettere altri codici con lo stesso metodo, senza limiti di numero.

La presenza di un codice restringe il ventaglio dei dati da immettere, come se i valori fossero stati designati come LEGALI. Si tenga conto, infine, che il programma ENTER non fa differenza fra lettere maiuscole e minuscole dei campi codificati; così "Epatite B", "EPATITE B" e "epatite b", nel campo MALATTIA, sono trattati tutti alla stessa maniera.

Se durante l'introduzione dei dati si ha la necessità di introdurre uno che non è né permesso né previsto dal codice, si può registrare lo stesso il dato usando però il tasto <Freccia Giù> invece di <Invio>.

I codici possono anche essere usati per dare all'operatore la possibilità di usare diversi sinonimi per riferirsi alla stessa cosa. Per utilizzare i codici in questo modo NON si deve collegare il campo ad un altro, in modo che il campo rimanga collegato a se stesso. Si introducono allora i sinonimi, premendo ogni volta il tasto <F8>. Per esempio, "Epatite B", "Ep B" o "B" possono essere tutti sinonimi che introdotti nel campo danno come risultato "EPATITE B".

I codici possono essere rimossi premendo <Ctrl-F8> sul codice da rimuovere e con il cursore su uno dei due campi collegati. Per far apparire tutti i codici presenti in un campo si preme <Maiusc-F8>.

COMANDO PER LA MODIFICA DEL CAMPO <F9>

I file .CHK non sono altro che serie di istruzioni da inviare al programma ENTER: è possibile allora redarre tali istruzioni indipendentemente dalle operazioni descritte in precedenza. Il comando per la redazione del campo, <F9>, permette di vedere, creare e modificare le istruzioni che normalmente sono inserite nel file dal programma CHECK, come se si lavorasse con un editore di testi. Questo lavoro può essere svolto però su un campo alla volta, quello in cui si trova il cursore.

Si provi a immettere qualche comando di CHECK in un campo: premendo il comando <F9>, con il cursore su quel campo, si può esaminare il risultato di quei comandi sul file .CHK che sarà letto da ENTER. Nell'esempio dato in precedenza sui codici per il campo MALATTIA, con <F9> appariranno i vari codici immessi e li si potrà modificare o aggiungerne direttamente sul file .CHK.

Si preme <F9> con il cursore sul campo MALATTIA. Apparirà una finestra, del tipo di quelle per la redazione di testi, con la porzione del file .CHK relativa al campo MALATTIA. Nella sezione "Codes" (Codici) si potranno notare delle coppie di codici e

alla fine la parola END (Fine). La prima di queste coppie dovrebbe essere "Epatite A" 11, ma si possono continuare a scrivere altre coppie avendo cura di racchiudere fra virgolette i codici che contengono spazi. Per cancellare un'intera riga si tiene premuto il tasto <Ctrl> e si digita la lettera Y, mentre per uscire dalla finestra bisogna premere <F10>. Rispetto al processo di codifica automatica, nella redazione diretta del file è più facile compiere errori. Ma se si sta attenti a non distruggere la struttura del file .CHK sarà molto più facile introdurre grandi gruppi di codici, come i nomi di città e regioni o una serie di malattie. Con questa procedura è infatti possibile introdurre e gestire parecchie centinaia di codici.

CAMBIARE I NOMI DELLE VARIABILI

I nomi delle variabili che Epi Info costruisce automaticamente in base al testo che precede gli spazi vuoti dei questionari, non sempre sono la rappresentazione migliore del significato della variabile. Questi nomi possono essere cambiati nel file di dati mediante l'editore del programma CHECK. Per cambiare il nome di una variabile si deve posizionare il cursore sul campo desiderato e premere <F9>. Premere dunque <F1>, che sta per "Change name" (Cambio del nome), come indicato in fondo allo schermo, e scrivere il nuovo nome del campo. Il nome della variabile sarà modificato sia nel file .CHK sia in quello .REC.

INTRODUZIONE DI COMMENTI AI "VALORI PERMESSI"

A volte è utile poter mostrare sullo schermo dei codici con la descrizione del relativo significato ma senza che siano registrati automaticamente nel file. Usando la funzione di redazione dei campi <F9>, ogni valore LEGAL può essere trasformato in un COMMENT LEGAL (valori permessi con commento) contenente la descrizione del valore in oggetto. Supponiamo che una serie di valori ammessi in un campo, in un file .CHK, sia come segue:

```
SESSO
  Legal
      M
      F
      S
  End
END
```

Premendo <F9> dalla schermo principale di CHECK quando il cursore è su SESSO, appariranno i valori come mostrati sopra. Aggiungere ora la parola COMMENT e le relative descrizioni ai codici:

```

SESSO
  Comment Legal
    M Maschio
    F Femmina
    S Sconosciuto
  END
END

```

Durante l'inserimento dei dati il programma ENTER terrà i codici e i commenti in una finestra nascosta che apparirà premendo il tasto <F9> (mostra i codici). Ciò è utile quando è presente un grande numero di codici e non si desidera avere sullo schermo tutte le loro descrizioni, come invece avverrebbe facendo introdurre i codici nel file direttamente dal programma CHECK.

Il rientro delle righe e l'incolonnamento esatto usato nell'esempio sono utili all'operatore per la chiarezza delle descrizioni e la lettura del file. Il programma Enter non ne tiene comunque conto e, se lo si desidera, si può ordinare il file in altro modo.

RICERCA AUTOMATICA

C'è un comando chiamato AUTOSEARCH (ricerca automatica) che esegue una ricerca di record doppi non appena si immettono i dati del questionario. Per esempio, si potrebbe dare l'istruzione di ricercare i record doppi, dopo aver introdotto il nome e il cognome di un paziente: immediatamente si verrà avvertiti della presenza di uno o più record che contengono già quel nominativo, evitando così di introdurre due volte i dati dello stesso questionario.

I FILE DI AIUTO PER I CAMPI DI ENTER

Un cenno soltanto a una nuova caratteristica di CHECK introdotta con la versione 6 di Epi Info: la possibilità di inserire, per ogni campo di ENTER, un testo di istruzioni per l'operatore o di spiegazioni sul campo. Per utilizzare questa possibilità è necessaria una certa esperienza; bisogna scrivere un file di guida e compilarlo, con opportune codifiche, con HELPPREP.EXE, fornito con Epi Info. Per i particolari leggere, nel capitolo 22 del manuale elettronico, il paragrafo sul sistema di menu.

LABORATORIO: CHECK

Avviamo Check dal menù principale, introduciamo il nome del file CORSO.REC, nello spazio "data file", e rispondiamo alla domanda "Ready?" (Pronto?) con Yes.

Check presenta il questionario sullo schermo, analogamente a ENTER. In basso troviamo tre righe con le indicazioni dei tasti funzione. Appena sopra questa serie di tasti funzione è mostrato il nome della variabile e il tipo di codice attivo nel campo sul quale si trova il cursore. I tasti funzione agiscono sul campo in cui è posizionato il cursore. Si scrive la condizione nello spazio della risposta. Premendo una volta un tasto funzione, si pone la condizione per quel campo e, in alcuni casi, premendolo una seconda volta la condizione viene tolta.

Collegamento di due campi

Nel nostro questionario di “partecipazione al corso” vogliamo legare il campo USL, nel quale abbiamo predisposto durante la creazione due tipi di campi, uno numerico e l'altro di testo. Ci posizioniamo con il cursore nel campo numerico, premiamo F5, il programma ci chiede quale campo deve essere legato a quello attuale, noi ci spostiamo sul campo da legare (a fianco) e ridigiamo F5.

In questo modo i campi sono collegati fra di loro.

Ora andiamo a codificare le USL.

Ci posizioniamo con il cursore nel primo campo, quello numerico scriviamo il n° della USL, la “110” ad esempio, premiamo F8, (Codes = codici), il cursore si sposta al campo successivo; qui scriviamo in chiaro il nome della USL (Ravenna).

Premendo di nuovo F8 il programma si ricorderà il legame tra il codice 110 e la parola “Ravenna”, e via di seguito.

Questa è una codifica, ad un nome corrisponde un codice che può essere un numero, una lettera dell'alfabeto o un segno qualsiasi, che ha la proprietà di “**stare per**” qualcos'altro.

In questo modo quando andremo ad inserire i dati e ci troveremo sul campo USL avremo la possibilità di aprire una finestrella dove ci sarà l'elenco di tutte le USL che avremo codificato, potremo scegliere quella che ci interessa e questa verrà scritta automaticamente nello spazio di risposta, oppure posizionandoci sul campo numerico e scrivendo il numero di USL che intendiamo inserire il programma scriverà automaticamente a fianco, in chiaro, il nome della USL al quale appartiene quel numero.

Anche ora, dall'interno di CHECK, possiamo visualizzare le finestrelle che vedremo in ENTER, premendo Maiusc + F5.

Valori permessi (legal)

Nella variabile sesso metteremo come valori permessi M e F.

A fianco del nome della variabile, nella quart'ultima riga del video, compariranno tutti i valori permessi.

Con **Maiusc + F6** si possono visualizzare in una finestrella.

Con **Ctrl + F6** si possono cancellare.

Capitolo 8

ANALYSIS: LISTE, FREQUENZE, TABELLE, STATISTICHE E GRAFICI

Se ENTER è il programma che serve per l'aggiornamento e la modifica delle basi di dati, ANALYSIS è il programma per *interrogare* la base di dati, cioè per analizzare i dati e ottenere le risposte per le quali si è costruita la base di dati.

ANALYSIS produce elenchi, frequenze, tavole, statistiche e grafici a partire dai file di Epi Info o di dBASE. Adoperando alcuni semplici comandi ANALYSIS seleziona record corrispondenti a determinati requisiti, ordina o elenca i record, calcola le frequenze e costruisce tavole di contingenza. Inoltre può compiere operazioni logiche e matematiche su una variabile, trasferire il risultato su una nuova variabile e indirizzare i risultati sullo schermo, alla stampante o in un file su disco.

I comandi costituiscono un mini-linguaggio di programmazione: possono essere digitati uno ad uno dalla tastiera, ma li si può anche inserire in un file che potrà essere avviato da ANALYSIS. In questo capitolo saranno descritte soltanto le operazioni più semplici, quelle che potrebbero essere usate nel corso di un'indagine su una serie di campi per condurre un'analisi preliminare del questionario.

PER CONOSCERE ANALYSIS

Avviando ANALYSIS dal menu di Epi Info si vedrà apparire sullo schermo una finestra inferiore, per inserire i comandi, e una finestra superiore più grande dove appaiono i risultati dell'esecuzione dei comandi. Sul margine inferiore dello schermo, alla base, sono indicati alcuni comandi, mentre in alto si trovano alcune informazioni sulla memoria disponibile e sui file di dati.

Ogni volta che il cursore è sulla scritta EPI6> si possono digitare i comandi dalla tastiera. Se si compie un errore o si vuole cambiare parte della linea del comando, si usano i tasti usuali per tornare sui caratteri scorretti e correggere. I tasti <Canc> e <Ins> funzionano come in EPED, cancellando i caratteri e alternando il modo Inserimento. I tasti <Inizio> e <Fine> mandano il cursore all'inizio o alla fine della riga sulla quale si trova.

I tasti funzione in basso sullo schermo permettono di selezionare degli argomenti di guida, dei comandi e una lista di nomi di variabili che appare premendo l'apposito tasto. Dopo aver premuto un tasto funzione si sceglie un'opzione dalla lista, muovendo l'evidenziatore in alto e in basso con le frecce, e si digita <Invio>.

Quando sulla finestra superiore appaiono i risultati delle operazioni effettuate, quelli già presenti vengono spostati in alto fuori dello schermo, ma possono essere riportati nella finestra usando i tasti <Pag Su> e <Pag Giù> e scorrere i risultati uno schermo per volta. Per muoversi in maniera più precisa si possono usare questi ultimi tasti in combinazione con il tasto <Ctrl>, con l'effetto di far muovere il contenuto della finestra una riga per volta.

LEZIONI

Nel menu DIDATTICA di Epi Info 6 ci sono due programmi interattivi per vedere all'opera le funzioni fondamentali di ANALYSIS. Si consiglia l'esecuzione dei due programmi, "Analisi dei dati parte 1" e "Analisi dei dati parte 2", anche prima della lettura di questo capitolo.

IL SISTEMA DELLE GUIDE (HELP)

Per vedere la lista dei comandi di ANALYSIS si preme <F2>, il tasto dei comandi, e sullo schermo apparirà una finestra con i nomi dei comandi disponibili. Selezionandone uno con <Invio> il comando sarà trasferito a fianco della scritta EPI6>, pronto per essere eseguito. Naturalmente il comando può essere digitato direttamente dalla tastiera. Una volta eseguito un comando, per esempio READ, si possono ottenere ulteriori informazioni sul comando stesso premendo <F1>, incluse quelle sul formato con cui devono essere forniti i nomi dei campi e gli altri dati. Se è stato inserito il comando, ma non è ancora stato eseguito con <Invio>, premendo <F1> apparirà una lista dei possibili argomenti di quel comando.

Per esempio con READ, dopo aver scritto il comando, premendo <F1> appare una finestra contenente tutti i file .REC essere letti (read). Basta quindi selezionarne uno e premere <Invio>. Le guide sono sempre a un passo da qualsiasi comando. In breve:

- Per informazioni generali su un argomento premere <F1>, senza il nome del comando sulla riga di EPI6>, e scegliere l'argomento dalla lista.
- Per informazioni sui comandi, premere <F2> e selezionarli.
- Per ulteriori informazioni sul comando, premere <F1> dopo averlo selezionato.
- Per informazioni sui file disponibili digitare READ, <Invio> e scegliere il file sul menu che appare.
- Per una lista delle variabili del file corrente, premere <F3>, selezionarle dalla lista e premere <Invio>. Si può scegliere più di una variabile: premendo il tasto "+" su un nome di variabile, questa è selezionata ("tagged"); mentre con il tasto "-" a selezione della variabile viene eliminata ("untagged").

Il sistema della guida è abbastanza intelligente da cambiare l'ordine delle parole in un comando; se dai menu si sceglie prima una variabile e poi un comando, il loro ordine sarà invertito automaticamente.

IL PRIMO PASSO: LEGGERE UN FILE

Un'analisi deve essere compiuta sui record di un file. Il file può essere stato prodotto sia da Epi Info, introducendo i dati con il programma ENTER, sia da dBASE. I file Epi Info da analizzare possono essere prodotti anche importando i dati da molti altri programmi, mediante il programma IMPORT.

Il comando che dice ad ANALYSIS quale file utilizzare è il comando READ <nome file>, ed è generalmente il primo comando che si dà in ANALYSIS.

Per vedere una lista dei file disponibili digitare:

READ

e premere <Invio>. Per vedere i file di un'altra directory, per esempio C:\Dati\, si dovrà scrivere **READ C:\Dati**, prima di premere <Invio>. Per vedere tutti i file di **dBASE**, scrivere **READ *.DBF**, e apparirà una lista di file di quel formato. Spostare l'evidenziatore con le frecce e scegliere un file premendo <Invio>.

Sia scegliendo i file da una lista, sia scrivendo **READ** e il nome del file, ANALYSIS userà questo stesso file per tutte le operazioni che seguiranno, fino a che non è introdotto un altro comando **READ**: questo file diventa l' "active data set" (il gruppo di dati attivo).

ANALYSIS può leggere e analizzare direttamente anche i file di **dBASE**. Per usare un file **dBASE** si scrive il nome del file aggiungendo semplicemente l'estensione .DBF. La sola operazione di ANALYSIS che non può essere eseguita su un file **dBASE** è l'aggiornamento (UPDATE) dei record, che sarà descritto in seguito.

I seguenti sono esempi di comandi READ corretti:

READ OSWEGO (l'estensione .REC è superflua)

READ OSWEGO.REC

READ OSWEGO.DBF (un file dBASE)

READ OSWEGO1 OSWEGO2 OSWEGO3

(Questi tre file, di formato identico, possono essere analizzati come un insieme unico)

Per il seguito della lezione si analizzerà OSWEGO.REC, fornito insieme ad Epi Info. Per selezionarlo si digiti:

EPI6>READ OSWEGO

ELENCARE I RECORD

Per elencare i record di un file, digitare: EPI6> **LIST**. Il risultato è il seguente:

REC	AGE	SEX	TIMESUPPER	ILL	ONSETDATE	ONSETTIME	BAKEDHAM	SPINACH	MASHEDPOTA
riga	età	sexo	ora pasto	malato	data insorg.	ora insorg.	prosciutto	spinaci	purè
1	11	M	9999	N		0	N	N	N
2	52	F	1500	Y	04/19	30	Y	Y	Y
-	-	-	-	-	-	-	-	-	-

Eccetera

Il comando LIST mostrerà soltanto tante variabili quante ne permette la larghezza dello schermo. Se si desidera elencare tutte le variabili, usare il comando:

EPI6>**LIST ***

Il segno "*" è un'abbreviazione che sta per "tutti i campi". LIST seguito da uno o più nomi di variabili, elenca solo quelle variabili; ciò costituisce un modo comodo per avere un elenco di nomi e indirizzi di persone da intervistare, o per omettere dei nomi e degli indirizzi da un elenco di dati. Il seguente comando:

LIST * NOT NAME ADDRESS

produrrà un elenco di tutti i campi *eccetto* i campi NAME e ADDRESS (nome e indirizzo).

FREQUENZE

Il comando per le frequenze (FREQ) serve a contare i casi di ogni categoria, all'interno di una variabile specificata, e a dare la frequenza assoluta e relativa di ciascuna categoria. Il comando:

EPI6>FREQ SEX

produce il seguente risultato:

Sex	Freq	Percent	Cum.
F	44	58.7%	58.7%
M	31	41.3%	100.0%
Total	75	100.0%	

Per primo è dato il numero dei casi (frequenza) di ogni categoria, seguito dalla percentuale sul totale e dalla percentuale cumulativa. Se il comando STATISTICS è impostato a ON e i campi sono di tipo numerico, saranno presentati anche la somma, la significatività statistica e la deviazione standard.

Il comando FREQ * produrrà le frequenze per tutte le variabili del questionario in oggetto: un'ottima strada per iniziare l'analisi di un nuovo insieme di dati. FREQ seguito da una serie di nomi di variabili, separate da spazi, produrrà frequenze separate per ciascuna delle variabili elencate.

TAVOLE DI CONTINGENZA

Il comando TABLES conta i record i cui valori soddisfano due campi contemporaneamente, creando in questo nodo una tavola di contingenza (chiamata anche tavola a doppia entrata o tabella). Così, il comando EPI6>TABLES SEX ILL darà il seguente risultato:

SEX	ILL		Total
	+	-	
F	30	14	44
M	16	15	31
Total	46	29	75

Analisi della tavola singola

Odds Ratio 2.01
 Limiti di confidenza al 95% di Cornfield per OR 0.70 < OR < 5.85
 Rischio relativo di (ILL=+) per (SEX=F) 1.32
 Limiti di confidenza al 95% di Greenland e Robins per RR 0.89 < RR < 1.96
 (Biometrics 1985; 41: 55-68)

Ignora il rischio relativo per studi caso-controllo.

	<u>Chi-quadrato</u>	<u>valori-P</u>
Non corretto	2.11	0.14679237
Mantel-Haentzsel	2.08	0.14951418
Correzioni di Yates	1.46	0.22620244

Il rischio relativo è dato dal rapporto fra il primo fattore della prima riga, diviso il totale della riga, e il primo fattore della seconda riga, a sua volta diviso per il totale della se-

conda riga. Se designiamo con "a" e "b" il primo e il secondo fattore della prima riga rispettivamente, e con "c" e "d" i due fattori della seconda riga, la formula del rischio relativo è la seguente:

$$(a / a + b) / (c / c + d) ; \text{ nell'esempio sopra: } (30/30+14)/(16/16+15) = 1.32$$

Insieme al rischio relativo è presentato, nell'esempio, anche un'interpretazione. Si tenga presente che l'interpretazione del rischio relativo (risk ratio) dipende dall'orientamento della tavola e che non tutti i rischi relativi saranno interpretabili. Se i fattori di rischio sono sul lato sinistro della tavola (in questo caso il sesso), e la malattia in cima alla tavola, indicando prima la sua presenza (+) e poi l'assenza (-), il rischio relativo rappresenta il rischio di malattia per le persone con il primo fattore, rispetto a quelle con il secondo fattore. Come indicato nella nota alla tavola sopra riportata, il rischio relativo dovrebbe essere ignorato in uno studio caso-controllo, dal momento che è privo di significato.

ANALISI STRATIFICATE

Si possono produrre tavole stratificate, elencando più di due variabili nel comando TABLES. Le variabili dopo le prime due servono come base per dividere le tavole in livelli, o strati, uno per ogni combinazione di variabili dopo la seconda. Per avere un esempio rapido si digiti:

TABLES VANILLA ILL SEX

ciò produrrà tavole separate a seconda del sesso, e le due tavole saranno seguite da un'analisi stratificata Mantel-Haenszel. Il sommario delle combinazioni statistiche dei due strati appare come segue per VANILLA ILL SEX:

** Sommario di due tavole con margini non-zero **
N per insiemi w/ risultati discordanti = 75

SOMMARIO DEGLI ODD RATIO	
OR grezzo per insiemi con risultati discordanti	23.45
Odd Ratio ponderata di Mantel-Haenszel	37.64
Limiti di confidenza al 95% per M-H OR	7.19 < OR < 197.01
Maximum likelihood estimate di OR (MLE)	28.58
Exact 95% confidence limits for MLE	6.01 < OR < 207.50
Exact 95% Mid-P limits for MLE	6.92 < OR < 160.22
Probability of MLE >= 28.58 if population OR = 1.0	0.0000011

Impossibile determinare l'interazione a causa di della presenza di valori a zero

SOMMARIO DEI RISCHI RELATIVI (Ignora per studi caso-controllo)	
RR grezzo	5.57
Sommario RR di (ILL=+) per (VANILLA=+)	5.73
Limiti di confidenza al 95% per RR	1.99 < RR < 16.51

** Fine dell'analisi stratificata **

Il valore molto alto dell'Odd ratio indica una forte associazione fra il consumo di gelato alla vaniglia e l'occorrenza della malattia. L'odds ratio ponderato di Mantel-Haenszel, comprendente i risultati della stratificazione per sesso, è persino più alta. Entrambi i limiti di confidenza e il valore P molto piccolo, indicano che l'Odd ratio ponderato di

Mantel-Haenszel è significativamente lontano da 1.0, il valore indicante l'assenza di associazione.

Entrambi i rischi relativi, grezzo e stratificato, sono anch'essi significativamente più alti di 1.0. Il gelato alla vaniglia fu dunque individuato con certezza, attraverso prove statistiche e microbiologiche, come la causa di questa diffusione di casi di intossicazione alimentare da stafilococco a Oswego, New York.

Per indicare tutte le variabili nel comando TABLES, può essere usato un asterisco o stella (*). Così, TABLES ILL * produrrà tavole di stati di malattia in base a ciascuna variabile del questionario. Questo è un metodo veloce per ottenere un'analisi preliminare del questionario, comparando i gruppi due a due.

Il test di Rothman, quando è possibile, indica le differenze negli odds ratio fra strati, se il valore P è piccolo. Tali differenze suggeriscono che l'odds ratio ponderato di Mantel-Haenszel, benché sia valido come consuntivo di tutti gli strati, non necessariamente riflette il valore per ciascuno degli strati.

IL COMANDO MEANS (MEDIE)

Il Comando TABLES è utile per i dati che possono essere ridotti in categorie e contati. Per l'analisi di dati continui come l'altezza, il peso, l'età, ecc., è più adatto il comando MEANS. Il comando MEANS produce una tavola che presenta i dati continui, o ordinali, ed esegue le statistiche appropriate, come ANOVA, ecc.

Il comando MEANS richiede due informazioni: la variabile contenente i dati da analizzare e la variabile che indica come distinguere i gruppi. Il comando è:

MEANS [Variabile numerica da analizzare] [Variabile per riunire in gruppi]

Se si desidera che non appaiano le tavole dei valori, aggiungere **/N** al comando che indica "No tables".

Usando il file OSWEGO.REC, **MEANS AGE ILL** comparerà le età delle persone per le quali ILL="Y" con quelle per le quali ILL="N".

Usando l'esempio MEANS AGE ILL, si dovrebbero ottenere i seguenti risultati:

ILL	Obs	Total	Mean	Variance	Std Dev
+	46	1806	39.261	477.264	21.846
-	29	955	32.931	423709	20.584
Difference				6.330	

ILL	Minimum	25%ile	Median	Maximum	75%ile	Mode
+	3.000	17.000	38.500	59.000	77.000	15.000
-	7.000	14.000	35.000	50.000	69.000	11.000

ANOVA

(solo per dati normalmente distribuiti)

Il valore p è equivalente a quello del test t di Student, dal momento che ci sono solo due campioni.

Variation	SS	df	MS	F Statistics	p-value	t-value
Between	712.655	1	712.655	1.560	0.213137	1.249147
within	33340.732	73	456.722			
Total	34053.387	74				

Test di Bartlett per l'omogeneità della varianza
Chi quadrato di Bartlett = 0.119 grado di libertà = 1 valore p = 0.729786

Le varianze sono omogenee con il 95% di confidenza.
Se anche i campioni sono normalmente distribuiti si possono usare i risultati di ANOVA.

Test "Two-Sample" di Mann-Whitney o Wilcoxon
(test Kruskal-Wallis per due gruppi)

H di Kruskal-Wallis (equivalente al Chi quadrato) = 1.161
Grado di libertà = 1
valore p = 0.281226

A prima vista sembrano esserci più numeri di quanto sia desiderabile, ma i risultati principali possono essere compresi istantaneamente. Le medie dei due gruppi di età sono 39 e 33. Si osservi ora il valore p sotto i test ANOVA e Kruskal-Wallis. Poiché entrambi sono superiori a 0.05, possiamo concludere che la differenza di età media fra i gruppi di malati e non malati non è significativa. Il test di Bartlett aiuta a decidere quale dei due metodi usare ma, in questo caso, entrambi i test conducono alla stessa conclusione.

Nota sui valori mancanti

Nelle procedure MEANS, TABLES, e FREQ, i valori mancanti saranno ignorati se sono stati immessi premendo <Invio> nel programma ENTER e se SET IGNORE (imposta "ignora") è in ON (normalmente lo è; si veda il comando SET nel capitolo 36 del Manuale elettronico). Se, tuttavia, si è impostato un altro codice per i valori mancanti, per esempio 99, ci si assicuri, durante la procedura per le medie statistiche, di selezionare solo valori non mancanti. Questo può essere fatto usando il comando, per esempio: SELECT AGE < > 99 (seleziona le età minori o maggiori di 99). Si faccia particolare attenzione su questo punto, specialmente se i dati sono stati importati da un altro programma in cui i valori mancanti sono codificati in maniera differente.

TITOLI

Il comando TITLE (titolo) permette di introdurre fino a cinque righe di testo, che appariranno in cima a una tabella, a una frequenza o a un grafico. TITLE può essere usato prima dei comandi TABLES, FREQ e i comandi per i grafici, in modo da produrre un titolo appropriato per presentare i relativi risultati. Per cancellare un titolo dopo che i suddetti comandi sono stati eseguiti, si può usare di nuovo TITLE lasciando lo schermo in bianco, oppure sostituendo altro testo. I seguenti comandi servono a definire un titolo a due righe da usare per altri comandi:

Title 1 Numero di casi in base alla data di insorgenza

Title 2 Bologna, Gennaio 1992

Per cancellare i titoli prima di eseguire altri comandi digitare:

Title 1

il titolo numero 1 sarà cancellato, così come tutti quelli di livello inferiore.

GRAFICI

ANALYSIS produce vari tipi di grafici direttamente a partire dai file di dati: istogrammi (HISTOGRAM), grafici a punti (SCATTER), grafici a torta (PIE), a linea spezzata (LINE) e a barre (BAR). Tracciare un grafico richiede un unico comando; le seguenti strutture di comandi possono darne un'idea:

HISTOGRAM <nome della variabile> [/Y]

PIE <nome della variabile>

BAR <nome della variabile> [/Y]

LINE <nome della variabile/i> [/X] [/Y]

SCATTER <nome della prima variabile> <nome della seconda variabile> [/R]
[/X][Y]

A volte è importante confrontare visivamente due grafici e, in tal caso, è importante che le coordinate del grafico (assi X, orizzontale, e Y, verticale) riportino delle scale di valori identiche. Per impostare i valori delle coordinate si aggiunge al comando il carattere barra obliqua in avanti ("/") e si specificano i valori minimi e massimi che riportano l'asse X e l'asse Y (il valore minimo deve essere sempre zero). Per esempio, il comando:

BAR ROLLS/Y=0-75

produce un grafico a barre, sulla variabile ROLLS (panini) in OSWEGO, il cui valore massimo riportato sull'asse Y è 75.

Per i grafici a punti (SCATTER), i numeri massimi e minimi degli assi X e Y possono essere ingranditi specificando nuovi numeri (per i grafici a punti le variabili sono necessariamente due), per esempio:

SCATTER AGE BP /X=0-300/Y=0-100

Se al comando SCATTER si aggiunge: /R, lungo i punti del grafico sarà tracciata la retta di regressione lineare. Si veda l'esempio seguente:

SCATTER SYSBP BODYMASS /R

in questo caso però, il minimo non può essere incrementato, né il massimo diminuito, rispetto ai valori scelti automaticamente.

STAMPARE I RISULTATI O TRASFERIRLI IN UN FILE

Prima di usare i comandi per la stampa, assicurarsi che la stampante sia collegata al computer, che sia accesa e fornita di carta. Premendo il tasto <F5> (Printer On), tutti i risultati che si vanno ad ottenere vengono trasferiti alla stampante, esclusi però quelli già presenti sullo schermo. Il modo migliore per stampare ciò che appare sullo schermo è premere <F5> (Printer On), poi il tasto freccia in alto per richiamare il comando precedente, cioè quello che ha prodotto il risultato sullo schermo (Analysis "ricorda" i comandi eseguiti in precedenza). Premendo <Invio> il comando sarà eseguito di nuovo, ma questa volta i risultati saranno stampati e non appariranno sullo schermo. Per escludere la stampante, alla fine, premere <F5> di nuovo (Printer Off).

Si noti che quando si preme <F5> appare: ROUTE PRINTER o ROUTE SCREEN. <F5> è in realtà una scorciatoia per questi comandi. Se si vuole, dunque, si possono scrivere quei comandi direttamente dalla tastiera.

Il comando ROUTE serve anche per trasferire i risultati in un file di nome qualsiasi. Per esempio:

ROUTE C:RISULTA.TXT

trasferirà i risultati in un file di nome RISULTA, con estensione .TXT (testo), nel drive C:. Questo fino a quando si esegue un altro comando ROUTE, per trasferire i risultati sullo schermo (SCREEN), alla stampante (PRINTER), o a un altro file; oppure se si chiude il file uscendo da ANALYSIS con <F10>.

Quando si dà il comando per un grafico, come descritto nella sezione precedente, e la stampante è attiva (Printer On), il grafico sarà stampato; altrimenti apparirà soltanto sullo schermo. La stampa dei grafici può essere eseguita con stampanti Hewlett-Packard Laserjet, altre stampanti e plotter compatibili, e su stampanti ad aghi IBM e EPSON compatibili, anche se altre stampanti di solito funzionano bene con liste e tavole. I grafici possono essere stampati anche con altre stampanti, ma la grafica è più soggetta a problemi di compatibilità rispetto al solo testo. Si faccia riferimento al comando SET, nel capitolo 36 del Manuale elettronico, per i metodi di impostazione di ANALYSIS con plotter e stampanti laser.

RICHIAMARE I COMANDI DI ANALYSIS

Non appena eseguito, ogni comando è memorizzato da ANALYSIS per una possibile riutilizzazione. Premendo il tasto <Freccia Su>, riappare il comando precedente a fianco alla scritta EPI6>, in modo da poter essere corretto o modificato. Una volta richiamata e redatta la riga, premendo <Invio> il comando sarà eseguito come se fosse stato appena scritto.

I comandi precedenti sono memorizzati in un settore chiamato "Command Stack". ANALYSIS salva di solito gli ultimi venti comandi, ma si può aumentarne il numero reimpostando la dimensione del settore "Command Stack". Si consulti, per questo, il comando SET nel capitolo 36 del Manuale Elettronico.

LABORATORIO: ANALIZZIAMO IL FILE ORIENTE.REC

Il file Oriente.Rec è un file di dati di esempio (presente nel quarto dischetto allegato a questo manuale) contenente la raccolta di informazioni su un episodio di tossinfezione alimentare verificatosi nella provincia di Ravenna.

Avviando Analysis dal menù principale ci troviamo di fronte a due finestre una **superiore** più larga dove troviamo scritto "Output Screen" (risultati), una inferiore più stretta, "Commands" (comandi), nella quale troviamo la scritta EPI6> ed il cursore che lampeggia: è da qui che verranno digitati i comandi, premendo <Invio> dopo aver introdotto il comando. Nella finestra superiore appariranno i risultati.

Digitando READ e premendo INVIO apparirà un menù con l'elenco dei FILE .REC che possono essere letti e che si trovano nel cassetto di lavoro, da dove abbiamo avviato Epi Info. Da questa finestra scegliere ORIENTE.REC con le frecce direzionali e premere

<Invio>. Per avere la conferma che il nostro File sia stato caricato, a sinistra in alto dopo il DATASET dovrà apparire il nome del File e tra parentesi il numero dei record contenuti.

Eseguiamo ora i seguenti comandi e osserviamo i risultati a video:

FREQ MALATO

Ci dà una tabella, dove troviamo sulla sinistra i valori che Analysis ha trovato in corrispondenza del campo indicato. Nel caso di MALATO, campo logico, Analysis ha trovato (+) cioè malati=si e (-) cioè malati=no. Sulla destra vi sono tre colonne: la prima indica la frequenza dei SI, quella dei NO, con il totale complessivo. La seconda mostra le percentuali sul totale. La terza con in cima CUM, che sta per cumulativa, rappresenta la somma cumulata delle percentuali.

FREQ ETA

Ci fornisce le età dei partecipanti al banchetto, la frequenza degli ospiti (evento = record = ospite) che aveva quella determinata età e la loro percentuale. In fondo alla tabella Analysis fornisce, solo per le variabili numeriche, alcuni indici sintetici quali la somma, la media e la deviazione standard dei valori trovati nel campo (in questo caso è interessante l'età media dei partecipanti)

TABLES SESSO PS

Ci troviamo davanti ad una tabella dove da una parte, a sinistra, abbiamo i valori della variabile sesso (M e F), e a destra la variabile Pronto Soccorso (PS), con (+) e (-) essendo un campo Y/N. Dentro la tabella 2x2 Analysis ha contato quanti maschi sono stati al PS, quante femmine, il totale dei ricorsi al PS e il totale generale.

Di seguito alle tabelle 2x2 Analysis produce alcuni indici epidemiologici, in automatico, RR = Rischio Relativo, O.R. = Odds Ratio e altro.

TITLE 1 Epidemia di tossinfezione a Milano Marittima

TITLE 2 Ravenna agosto 1991.

TITLE 3 Frequenza della malattia

FREQ MALATO

In tabelle o prodotti successivi si può cambiare il titolo e adattarlo al tipo di elaborato che seguirà. È sufficiente riscrivere il comando con il testo aggiornato. Il comando TITLE 1 senza specificare alcun titolo, elimina tutti i titoli.

LINE DURATASINT

Produce un grafico a linee. Come si vede sull'asse delle X, quello orizzontale, sono ordinati i valori che Analysis trova nella variabile specificata: la durata in ore dei sintomi. Sull'asse delle Y, quello verticale c'è il numero dei casi di ciascuna categoria, quindi la frequenza.

HISTOGRAM ORAINSORGE

Questa è una curva epidemica.

Per ulteriori dettagli su Analysis si vedano i capitoli 9 e 10 di questo volume e i capitoli 9, 11, 13, 32, 36 del Manuale elettronico.

Capitolo 9

SCRIVERE PROGRAMMI E IMPOSTARE I DATI PER ANALYSIS

INTRODUZIONE

Nel capitolo precedente si sono presentati i metodi per produrre in ANALYSIS risultati statistici, tavole e grafici, a partire dai file di dati. Per insiemi di dati già perfettamente "sgrossati" questo sarebbe sufficiente per fare un'analisi completa. Le situazioni reali, tuttavia, producono dati spuri e spesso la maggior parte del tempo necessario a un'analisi serve solo per sgrossare i dati.

Questo capitolo descrive i metodi per visionare e modificare i dati in ANALYSIS: ordinare i record in ordine alfabetico o numerico, includere o escludere gruppi selezionati di record, definire e assegnare valori a nuove variabili, riunire o ricodificare le variabili, eseguire operazioni in base a condizioni (operatore IF), manipolare date e intervalli temporali e infine impostare una serie di parametri di ANALYSIS per modificarne il funzionamento.

Queste operazioni non sono difficili, prese individualmente, ma un determinato insieme di dati può richiederne molte di seguito, prima che tutti i campi siano pronti in maniera soddisfacente per l'analisi. È utile incorporare in un file di programma i comandi per tali operazioni, in modo che il programma possa essere svolto diverse volte dopo aggiunte e aggiustamenti.

SCRIVERE PROGRAMMI

I programmi sono file creati con un editore di testi e contenente gli stessi comandi di ANALYSIS che si useranno. Tali file possono essere creati con EPED o con altri editor di testi, purché producano file di testo (.TXT). A questo scopo, in EPED si deve impostare a OFF il Word Wrap e in Wordstar, per esempio, si deve usare il modo Non-documento.

Creando un file di programma di nome SAMPLE.PGM (file campione), il comando RUN (avvia) potrà essere usato per eseguire tutti i comandi contenuti in questo file:

```
EPI6>RUN SAMPLE.PGM
```

L'estensione .PGM è facoltativa, ma è un buon metodo usare questa distinzione per i file di programma.

Il seguente programma, creato con EPED, legge il nostro solito file OSWEGO.REC, seleziona i record con casi di insorgenza di malattia (ONSETDATE) al solo giorno 18 e compie le seguenti operazioni: ordina questi record in base a SEX e AGE; indirizza in un file di testo, di nome RESULTS.TXT, una lista di variabili selezionate da questi record; indirizza in un file ASCII, di nome RESULTS.TAB, una tavola 2x2 di SEX e AGE; trasmette tutti i risultati sullo schermo (si noti l'uso dei comandi TITLE) e avverte l'operatore che la procedura è terminata.

```

* Elenchi e tavole di contingenza di OSWEGO.REC
* Solo il giorno 18 del mese
* Gli asterischi servono a nascondere questi
READ OSWEGO [Leggi il file OSWEGO]
SELECT ONSETDATE="04/18" [Seleziona le insorgenze del 18/4]
SORT SEX AGE [Ordina per sesso ed età]
ERASE RESULTS.TXT [Cancella il file di testo RISULTATI]
ROUTE RESULTS.TXT [Indirizza al nuovo file di testo RISULTATI]
LIST SEX AGE ILL [Elenca i campi SESSO, ETÀ, MALATI]
ERASE RESULTS.TAB [Cancella il file di tabella RISULTATI]
ROUTE RESULTS.TAB [Indirizza al nuovo file di tabella RISULTATI]
TITLE 1 CHURCH SUPPER QUESTIONNAIRE - OSWEGO
[Titolo 1: Questionario per il pasto parrocchiale di OSWEGO]
TITLE 2 AGE X SEX [Titolo 2: Tabella ETÀ per SESSO]
TITLE 3 CROSSTABS:DATE=18 [Titolo 3: Tabella per il giorno 18]
TITLE 4 CLASS EXAMPLES [Titolo 4: Classe di esempi]
TABLE AGE SEX [Produce la tabella con entrate: ETÀ, SESSO]
TITLE 1 TABLE 2 - OSWEGO [Titolo 1 : Tabella 2 - OSWEGO]
TITLE 2 AGE X ILL [Titolo 2 : Tabella ETÀ per MALATTIA]
TABLES AGE ILL [Produci la tabella con entrate : ETÀ MALATTIA]
TITLE 1 TABLE 3 - OSWEGO [Titolo 1 : Tabella 3 - OSWEGO]
TITLE 2 [Titolo 2 : - vuoto, per cancellare il precedente-]
TABLES * ILL [Produce la tabella con * entrate per MALATTIA]
ROUTE SCREEN [Indirizza i risultati sullo schermo]
ECHO LA PROCEDURA È TERMINATA. PREMERE <F10> PER TORNARE AL MENU

```

I programmi possono essere corti fino a una sola riga oppure lunghi a piacere. Sono comunque molto utili quando le operazioni sono di una certa complessità e il compito di scriverle ripetutamente diventa pesante. I programmi possono essere costruiti per le persone che usano raramente Epi Info, oppure per gli operatori che ci lavorano quotidianamente in modo che, per esempio, la produzione di un rapporto settimanale richieda l'esecuzione del solo comando RUN WEEKLY.PGM.

I programmi sono anche molto utili per cercare di correggere gli errori di righe di comando contenenti operatori logici complessi, cioè molti IF, AND e OR di seguito (in italiano questi operatori sono: SE, E, O, rispettivamente). Si scrive, per provarlo, il programma contenente tali operatori, e poi lo si avvia diverse volte (con RUN) per vedere come funziona. Con questa procedura per prove ed errori sarà facile scoprire i difetti del programma, che potrà così essere riutilizzato in futuro senza possibilità di errore. Nei programmi è inoltre molto utile impostare i comandi READ, SELECT, RECODE, IF e LET per l'analisi preliminare di un file .REC. Avviato questo programma si potrà poi procedere all'analisi vera e propria, con i comandi per la produzione di elenchi, frequenze, tabelle e medie statistiche.

SALVARE E RICHIAMARE I COMANDI DI ANALYSIS

Una maniera comoda per costruire un programma è digitare i comandi dalla tastiera interagendo con il programma. Una volta compiuto un lavoro, si possono salvare in un file tutti i comandi che si sono eseguiti (fino al limite del "Command Stack", la memoria di transito, che di solito è di 20 righe) usando il comando SAVE insieme a un nome di file. Per esempio:

```
SAVE LAVORO.PGM
```

Tutti i comandi eseguiti per produrre quel lavoro, e contenuti nella memoria di transito, verranno trasferiti nel file, il quale potrà essere modificato con EPED per eliminare alcuni comandi o aggiungerne di nuovi. SAVE è molto utile quando, dopo aver eseguito una serie di comandi, si decide di farne un programma permanente. Con SAVE si salvano quei comandi in un file .PGM e con EPED si rifinirà il programma per adattarlo al suo uso futuro in ANALYSIS.

Se si sono salvati i comandi contenuti nella memoria di transito, e si desidera richiamarli dal file, lo si può fare con il comando LOAD, carica.

LOAD <nome del file>

ritrasferirà i comandi dal file alla memoria di transito attuale, dove potranno essere modificati. La differenza fra questo comando e il comando RUN è che i comandi ritrasferiti dal file non sono eseguiti: per farlo si dovrà selezionarli con i tasti freccia e premendo <Invio>.

La memoria di transito (Command Stack) può essere aumentata per contenere più comandi. Per fare questo si veda il Manuale elettronico, precisamente il comando SET CMDSTACK al capitolo 36.

ACCESSO AL DOS DIRETTAMENTE DA ANALYSIS

Si possono eseguire comandi DOS e perfino avviare altri programmi direttamente da ANALYSIS, senza uscire dal programma. Premendo il tasto <F9> appare la scritta DOS C:>, dalla quale si può eseguire qualsiasi comando DOS e far partire qualsiasi programma. Per tornare in ANALYSIS si digita **EXIT**, e ci si ritroverà allo stesso punto di prima con tutte le procedure in corso intatte. Mostriamo l'utilità dell'accesso al DOS con un esempio: la creazione di una copia di riserva (backup copy) di un file di dati. Supponiamo che il file sia OSWEGO.REC. Premere <F9> per accedere al DOS e digitare:
C:>COPY OSWEGO.REC OSWCOPIA.REC

Il comando DOS, COPY, eseguirà la copia di riserva del file OSWEGO.REC. Scrivere ora:

EXIT

istantaneamente riappare ANALYSIS pronto per continuare il lavoro in corso.

ORDINARE I RECORD: IL COMANDO SORT

Molte volte è importante che i record di un file siano ordinati in una maniera particolare, per esempio in ordine alfabetico usando il campo NAME, oppure in ordine di età, utilizzando il campo AGE. La creazione di un ordine è particolarmente importante in vista dei comandi LIST (per elenchi) e WRITE (per la scrittura). Il comando SORT (ordina) farà al caso, ordinando anche molti campi contemporaneamente, se si vuole. I record sono temporaneamente disposti nel nuovo ordine e possono essere elencati (con LIST) o scritti su un altro file con WRITE RECFILE.

Per ordinare alfabeticamente in base ai nominativi digitare:

SORT NOME

Per ordinare con criteri basati su più di un campo, porre per primo il campo di interesse principale e di seguito gli altri. Se per esempio si vuole ottenere un elenco contenente prima tutte le donne seguito da quello degli uomini, e fare in modo che ciascun gruppo sia in ordine di età, si usi il comando:

SORT SESSO ETA

Questo comando disporrà i record prima in base al sesso e, all'interno di ciascuna categoria, li ordinerà in base all'età. L'ordine che risulta dall'esecuzione di SORT scompare se si legge (comando READ) un altro file, oppure se si esce da ANALYSIS. Usando ancora i comandi ROUTE e WRITE RECFILE, si possono salvare i file già ordinati e rendere dunque quell'ordine permanente (si veda anche il paragrafo: creazione di un nuovo file Epi Info, nel Capitolo 13 del manuale elettronico). Non è comunque necessario ordinare i record per eseguire i comandi per le frequenze e per le tabelle, dal momento che in questi casi l'ordine è creato automaticamente dal programma.

SELEZIONARE I RECORD DA ANALIZZARE

ANALYSIS normalmente elabora tutti i record di un file. Si possono però selezionare dei gruppi di record che rispondono a determinati requisiti con il comando SELECT. I seguenti comandi servono alla selezione dei record che si riferiscono ai soli maschi di età compresa fra i 20 e i 49 anni:

```
EPI6>SELECT SEX = "M"
```

```
EPI6>SELECT AGE < 50
```

```
EPI6>SELECT AGE > 19
```

I comandi SELECT eseguiti in successione si sommano come se fosse stato introdotto l'operatore logico AND. Infatti i tre comandi esposti sopra potrebbero essere scritti in un'unica riga congiungendo con AND i tre requisiti:

```
EPI6>SELECT (SEX = "M") AND (AGE > 19) AND (AGE < 50)
```

si ricordi che una proposizione con l'operatore logico AND - congiunzione - è vera quando sono vere **tutte** le condizioni che sono congiunte; il comando selezionerà allora tutti i record per i quali sono veri il sesso maschile, l'età minore di 50 anni e maggiore di 19. Le parentesi sono necessarie per separare le condizioni ed è più sicuro usarle sempre per evitare errori. SELECT accetta, comunque, anche forme abbreviate di condizioni multiple:

```
EPI6>SELECT AGE <50 AND >19
```

Si faccia attenzione alla logica; i seguenti comandi:

```
EPI6>SELECT GROUP = 1
```

```
EPI6>SELECT GROUP = 2
```

non selezioneranno alcun record da analizzare. Infatti non c'è nessun record con il valore 1 per la variabile GROUP e **contemporaneamente** con il valore 2 nello stesso campo. Probabilmente il comando corretto è:

```
EPI6>SELECT (GROUP = 1) OR (GROUP = 2)
```

Questo comando selezionerà **sia** i record appartenenti al gruppo 1, **sia** quelli appartenenti al gruppo 2. L'operatore logico OR - disgiunzione - rende vera una proposizione

quando **almeno una** delle condizioni contenute è vera. La selezione avverrà dunque per i record per i quali è vero che hanno la variabile GROUP o con valore 1 o con valore 2.

SELECT può anche essere usato per escludere record. Se si vogliono eliminare i record appartenenti al gruppo 9, per esempio, si potrà usare il programma ENTER per cancellare i record uno ad uno. Con SELECT, invece, è sufficiente eseguire questo comando:

```
EPI6>SELECT GROUP < > 9
```

i segni "< >" significano o minore o maggiore, dunque: "non uguale a".

Le condizioni di selezione rimangono operanti fino a quando non si legge (comando READ) un altro file, si esce da ANALYSIS, oppure si cancellano tutte le condizioni eseguendo il comando:

```
EPI6>SELECT
```

senza inserire alcun requisito. Mediante i comandi ROUTE e WRITE RECFILE (scrivi un file .REC), infine, si possono trasferire in file permanenti i soli record selezionati.

ESAMINARE E MODIFICARE I RECORD CON BROWSE E UPDATE DI ANALYSIS

Ci sono diversi metodi per rivedere e modificare i dati. Il programma ENTER può essere usato per esaminare e anche per cambiare ciascuno dei record. In ANALYSIS si possono calcolare le frequenze di ogni campo, con il comando `FREQ *`, ed esaminare i dati che si discostano troppo dalla media per vedere se si tratta di un errore.

ANALYSIS offre però un altro metodo, con il quale i record sono mostrati sotto forma di un tabellone logico, come nei fogli elettronici. Usando il comando BROWSE (sfogliare), i record possono essere esaminati sotto questa forma. Il comando UPDATE (aggiorna) è eseguibile solo separatamente da BROWSE, in modo da non cambiare inavvertitamente i dati se si ha solo intenzione di sfogliarli. UPDATE permette di cambiare permanentemente i dati del file senza uscire da ANALYSIS. Dal momento che UPDATE modifica irreversibilmente il file di dati, prima di usarlo ci si assicuri di avere a disposizione una copia del file .REC in oggetto.

BROWSE (sfogliare)

Per usare BROWSE premere <F4>, il tasto per sfogliare. I record del file corrente appaiono sullo schermo sotto forma di foglio elettronico. Muovendo il cursore a destra o in basso, è possibile esaminare i campi di tutti i record. Se si preferisce il formato del questionario originale premere <F4> per avere un "Full Screen" (a pieno schermo), come indicato in basso sullo schermo. Il questionario apparirà come in ENTER e i record potranno essere "sfogliati" usando i tasti <F7> e <F8>. Per conoscere gli effetti degli altri comandi indicati in basso sullo schermo si consiglia di sperimentarli direttamente, senza bisogno di descrizioni. Sia il tasto <Esc> sia <F10> servono per abbandonare il programma e tornare alla scritta EPI6>.

UPDATE (aggiornare)

UPDATE funziona esattamente nella stessa maniera di BROWSE, solo che bisogna digitare UPDATE sulla riga dei comandi, invece di utilizzare un tasto funzione come per

BROWSE. Apparirà lo stesso foglio elettronico, ma il campo corrente è posizionato in cima al foglio, pronto per ricevere dati e correzioni. In UPDATE funzionano alcuni, ma non tutti, i comandi di CHECK. Quando si introduce una modifica nel campo corrente il record, e il campo contenente il cursore, vengono modificati permanentemente nel file di dati. Sicuramente si troverà UPDATE uno strumento comodo per esaminare i dati e per correggere gli errori di battitura o gli errori di codifica. Tuttavia UPDATE non può essere usato per modificare o aggiungere le variabili create da ANALYSIS: si possono aggiornare *soltanto i dati immessi*.

Dal momento che abbiamo appena fatto per sicurezza una copia del file OSWEGO.REC, possiamo aggiornarlo con UPDATE digitando i seguenti comandi:

READ COPIAOSW.REC

UPDATE

Quando si è finito di aggiornare il file con UPDATE, si digiti il comando:

ERASE COPIAOSW.REC

BROWSE e UPDATE possono mostrare dei sottoinsiemi del file di dati (gruppi di record selezionati dal tutto in base a determinati criteri). Il modo di selezionare (select) i record è descritto in un paragrafo precedente.

DEFINIRE NUOVE VARIABILI

Per molte operazioni di "pulizia" dei dati, come la ricodifica delle variabili, bisogna creare nuove variabili. I risultati della ricodificazione o di altre operazioni sono posizionati temporaneamente nelle nuove variabili, e appaiono sullo schermo o alla stampa. Queste variabili definite scompaiono quando si esce da ANALYSIS, ma possono essere conservate in un file permanente con il comando ROUTE <nome del file> e con WRITE RECFILE. I comandi che hanno creato le nuove variabili possono anch'essi essere conservati in un file di programma, come descritto all'inizio del capitolo, in modo da ricreare le stesse variabili avviando quel programma.

Le variabili definite possono contenere quattro tipi di dati: testo, numeri, date europee e date americane. Le variabili di testo supportano i caratteri alfabetici, incluse le cifre, e possono comprendere: "Mario", "ID4956", "Si", "(0544) 409099" e "N° risult. lab.". Le variabili numeriche devono contenere solo numeri, per esempio 3 o 40.90.99, ma non (0544), perché le parentesi non sono dati numerici. Le date hanno la forma mm/dd/yy oppure dd/mm/yy (mm/gg/aa; gg/mm/aa).

Per definire le nuove variabili in ANALYSIS si usano gli stessi caratteri speciali, e la stessa lunghezza, delle variabili usate nel questionario (si veda il Capitolo 4). Sono così permesse le seguenti forme di variabili:

DEFINE NOME _____

DEFINE PREZZO ###.###

DEFINE AGE ###

DEFINE INSORG <mm/dd/yy>

DEFINE ID #####

DEFINE INSORG <dd/mm/yy>

Non sono invece permessi i formati: <phonenum>, <IDNUM>, <Y>, <A> e <Today>. Si possono però introdurre queste variabili come variabili di testo.

Le variabili definite di tipo STANDARD, sono convertite in valori mancanti non appena ANALYSIS legge ciascun record del file, per esempio durante i comandi `FREQ`, `LIST` o `TABLES`.

Le variabili definite di tipo CUMULATIVE (additive) sono variabili i cui valori sono sommati l'uno all'altro, o accumulati altrimenti, dal primo all'ultimo record del file. Queste variabili sono azzerate mediante l'esecuzione del comando `READ` e dopo i comandi che leggono l'intero file, come `LIST`, `FREQ` o `TABLES`. Per trasformare una variabile in tipo CUMULATIVE, si scrive la parola "CUMULATIVE" dopo la definizione della variabile.

Le variabili definite di tipo GLOBAL si creano ponendo la parola GLOBAL dopo la definizione della variabile. Queste variabili hanno due proprietà che le distinguono dalle variabili di tipo standard:

1. Conservano la loro definizione anche quando è avviato un altro programma di ANALYSIS. Sono dunque utili per trasportare un messaggio da un programma ad un altro, durante una sessione di ANALYSIS, senza interruzioni. Le variabili standard, invece, scompaiono quando è avviato un nuovo programma.
2. Analogamente alle variabili CUMULATIVE conservano il loro valore alla lettura di ogni nuovo record.

Le variabili GLOBAL vanno impostate inizialmente come valori mancanti (missing). Tuttavia si può impostare qualsiasi valore iniziale desiderato, il quale va indicato dopo la definizione. Per esempio:

```
DEFINE DATA <dd/mm/yy> GLOBAL "03/03/90"
```

creerà una variabile DATA, di tipo GLOBAL, avente come valore iniziale la data europea 3 marzo 1990.

Nei prossimi due paragrafi si vedrà come si possono assegnare valori alle variabili definite e come si possono assegnare loro valori di variabili ricodificate di campi esistenti.

COME SONO ESEGUITI I COMANDI

Scrivendo i programmi e usando variabili definite è importante capire in che modo sono eseguiti i comandi. Alcuni comandi, come `LIST`, `FREQ`, `TABLES`, `MEANS`, `MATCH`, e i comandi per la grafica, fanno sì che ANALYSIS legga ogni record del file corrente per mandare poi i risultati sullo schermo o alla stampante. Sono elaborati soltanto i record selezionati al momento (con `SELECT` e con l'opzione "delete" (cancella) impostata dal comando `SET PROCESS`). I record riportanti valori mancanti nei campi in corso di elaborazione sono ignorati se l'opzione `SET IGNORE` è impostata a ON.

Alcuni comandi, come `LET`, `IF`, `RECODE` e `WRITE`, sono eseguiti ogni volta che il record è elaborato. Così, se il file contiene 3000 record, il comando `IF AGE >20 THEN AGELABEL = "ADULT" [SE ETÀ >20 ALLORA metti ETICHETTAETÀ = "ADULTO"]`, sarà processato 3000 volte durante il comando `TABLES`, ogni volta che ciascun record è inserito nella tabella. Di conseguenza la tabella mostrerà 1765 "ADULTI", esattamente come se il campo AGELABEL fosse stato presente nel questionario originale e un impiegato avesse introdotto nel campo 1765 volte la parola adulto.

Alcuni comandi, come IMMEDIATE, TYPE, ECHO e DEFINE, sono eseguiti una sola volta, nel punto dove sono posizionati nel programma. Questi comandi servono, di solito, per comunicare con l'utente attraverso lo schermo e sono descritti più diffusamente nel Manuale elettronico, al capitolo 13.

Un comando chiamato PROCESS consente ad ANALYSIS di leggere un file e di eseguire tutti i calcoli associati ad ogni record. Quindi sarà possibile impostare una variabile in modo che calcoli una somma cumulativa per un determinato campo, ad esempio:

```
DEFINE SOMMA ##### CUMULATIVE
```

```
SOMMA = SOMMA+COSTO
```

```
PROCESS
```

```
TYPE "@SOMMA"
```

PROCESS dà il consenso all'esecuzione del calcolo, e si potrà poi usare un comando TYPE per visualizzare i risultati dopo che tutti i record sono stati elaborati.

ASSEGNARE UN VALORE SINGOLO: IL COMANDO LET

Il comando LET, come i comandi RECODE e IF, è eseguito ogni volta che il record è calcolato. LET assegna un valore a una variabile definita, usualmente il risultato di un calcolo aritmetico.

L'operazione più semplice è assegnare il valore di una variabile esistente a una nuova variabile costruita da ANALYSIS. Se per esempio il questionario in corso di analisi contiene la variabile ETA (senza accento in quanto è un simbolo non permesso), il seguente comando definisce una nuova variabile, chiamata ETA2, e copia il valore di ogni record di ETA in quelli di ETA2:

```
DEFINE ETA2 ##
```

```
LET ETA2 = ETA
```

L'uso della parola "LET", letteralmente "fai" o "fa sì che", è facoltativa. Pertanto il secondo comando può essere scritto:

```
ETA2 = ETA
```

Il comando LET può essere usato in combinazione con qualsiasi operazione logica o matematica. Per una lista di queste operazioni si veda il Manuale elettronico in linea di Epi Info, sotto la voce OPERATORI del capitolo 36. Per esempio, per ottenere una variabile che rappresenti il quadrato del valore della variabile PESO esistente nel questionario, i comandi sono i seguenti:

```
DEFINE PESOQUADR #####
```

```
PESOQUADR = PESO*PESO [ "*" è il simbolo della moltiplicazione ]
```

Si faccia attenzione a porre sempre a destra del segno "=" le variabili già esistenti, e a sinistra quelle che si vanno a definire. È possibile usare una variabile definita di tipo "globale", per contare o accumulare i valori di record in record. Per esempio, se nel questionario non si è inclusa una variabile per il numero di record, la si può costruire usando LET nel modo seguente:

```
DEFINE NUMEROREC ##### CUMULATIVE
```

LET NUMEROREC = NUMEROREC + 1

In fase di elaborazione, alla nuova variabile NUMEROREC sarà assegnato il numero zero e, ad ogni nuovo record, sarà assegnato un numero di 1 maggiore del precedente. L'effetto sarà che i record saranno contati e numerati sequenzialmente a partire da 1.

RICODIFICARE I DATI

I dati introdotti con i questionari non sempre hanno la forma adatta per essere elaborati. Per esempio è desiderabile, a volte, dividere le età in gruppi di cinque anni quando nel questionario queste sono introdotte singolarmente. Oppure si vogliono dividere i record in base alle regioni di appartenenza degli intervistati, mentre nei record sono registrate solo le province, o ancora si desidera sostituire alcune descrizioni con dei numeri in codice. Il comando RECODE ("ricodifica") permette di assegnare nuovi valori ai dati esistenti e di introdurli in nuove variabili definite.

Per usare il comando RECODE bisogna innanzitutto definire una nuova variabile, adatta a ricevere i dati, con il comando DEFINE:

DEFINE GRUPPOETA STRING WIDTH=10

[campo di tipo testo di ampiezza = 10 caratteri, usato al posto di 10 caratteri di sottolineatura]

RECODE ETA TO GRUPPOETA BY 10

[ricodifica ETA su GRUPPOETA]

questo comando suddividerà i valori della variabile ETA in gruppi di valori di dieci anni, trasferendoli su GRUPPOETA, in modo da poter usare la nuova variabile GRUPPOETA in tabelle o calcoli di frequenza. È possibile però specificare in modo più preciso i codici di ogni gruppo definito. Per esempio:

DEFINE GRUPPOETA STRING WIDTH=11 [ampiezza campo = 11 caratteri]

RECODE ETA TO GRUPPOETA 1-12=BAMBINO 13-19=ADOLESC.

20-35=GIOVANE 36-65=ADULTO ELSE=ANZIANO

dividerà le età in gruppi i cui limiti sono fissati dall'utente. La parola inglese ELSE nel comando (letteralmente "altrimenti") può anche essere omessa. In questo caso i valori che non rientrano nei gruppi prefissati saranno considerati come valori mancanti. Ecco un altro esempio:

DEFINE NOME CITTÀ STRING [campo tipo testo]

RECODE CODCITTÀ TO NOME CITTÀ 1=BOLOGNA 2=RAVENNA

Il calcolo delle frequenze della variabile NOME CITTÀ, i cui valori (le città) sono codici, presenterà i nomi "RAVENNA" e "BOLOGNA" piuttosto che i relativi codici "1" e "2". Ulteriori dettagli sull'uso di RECODE si trovano nel Manuale elettronico, capitolo 36.

PORRE CONDIZIONI CON IF/THEN

Andiamo ancora una volta all'esempio di indagine per tossinfezione alimentare condotta a Oswego, i cui dati sono contenuti nel file OSWEGO.REC. Supponiamo di voler sapere quanti, fra malati e non malati, hanno consumato latticini. I prodotti contenenti latticini sono LATTE, CIOCCOLATA (in tazza) e VANIGLIA (gelato). Per prima cosa va definita una nuova variabile, per esempio LATTICINI:

```
DEFINE LATTICINI
```

```
LET LATTICINI = "N"
```

La seconda riga del comando assegna provvisoriamente un valore negativo alla variabile ("N" significa "No"). Quindi si pongono le condizioni, con IF/THEN ("se/allora"), per determinare quali casi possono tramutare il valore di LATTICINI in "S" (positivo). Per esempio:

```
IF (VANILLA = "Y") OR (MILK = "Y") OR (CHOCOLATE = "Y") \
THEN LATTICINI = "Y"
```

Questo comando significa che: *se* almeno una fra le variabili, MILK (LATTE), CHOCOLATE (CIOCCOLATA) o VANILLA (GELATO), ha valore positivo, *allora* la variabile LATTICINI assume un valore positivo, cioè il soggetto ha consumato un prodotto caseario. La ragione per cui si assegna prima il valore "N" a tutti i record per poi tramutarne solo alcuni in "Y" è che, altrimenti, dopo il primo valore "Y" assegnato, ANALYSIS assegnerebbe lo stesso valore a tutti i campi vuoti.

Si tenga presente che se la condizione è troppo lunga per stare in una sola linea, si deve scriverla con l'editore di testi EPED e usare la barra retroversa "\" (backslash) ad ogni fine riga, per avvertire ANALYSIS che si tratta di un'unica condizione. Con questa tecnica le condizioni possono essere lunghe fino a 254 caratteri.

Nella stesura delle condizioni è permesso, per facilità di lettura, racchiudere fra parentesi ogni singola condizione, o parte di essa, e spaziare le parole. Non è però obbligatorio farlo. I valori da assegnare alle variabili devono essere fra virgolette ("Y", "N"), mentre i nomi delle variabili vanno lasciati senza. Non sono permesse, invece, le cosiddette nidificazioni di condizioni, cioè condizioni che sottostanno a loro volta ad altre condizioni.

Le condizioni formulate con IF sono calcolate ogni volta che si processa un record. Così, se si ha un file contenente 3000 record, il comando:

```
IF ETA > 100 THEN TYPE "Errore: Età > 100"
```

farà apparire un messaggio di errore tante volte quanti sono i record contenenti un'età maggiore di 100. In alcuni casi, tuttavia, può essere comodo che un certo comando sia eseguito una sola volta. A questo proposito si usa il comando IMMEDIATE, descritto più dettagliatamente nel Manuale elettronico al capitolo 36.

DATE

Le date hanno sempre destato delle preoccupazioni, sia prima che dopo Papa Gregorio XIII che inventò il calendario moderno nel 1582. Bisogna avere a che fare con anni bisestili, con ordini europei o americani, con diversi numeri di giorni in ogni mese e con il fatto che se l'anno è divisibile per 100 non è bisestile, a meno che non sia divisibile anche

per 400. Senza poi contare il fatto che, sembra, l'anno che consideriamo come il primo sia in realtà compreso fra il 6 e il 4 avanti Cristo. Considerate queste premesse Epi Info fa il suo meglio per trattare le date in modo agevole e comodo. Le variabili per le date possono essere create sia direttamente dal questionario originale, sia da ANALYSIS come nuove variabili definite. I tipi di variabili ordinari per le date da inserire nel questionario sono:

<mm/gg> formato americano
<mm/gg/aa>
<gg/mm> formato europeo e militare
<gg/mm/aa>
<today> data del giorno
<today/aa>
<today/aaaa>

Per ordinare i record usando le date, si usa il comando SORT. Epi Info si occuperà di riordinare i record in ordine cronologico, ponendo le date più antiche all'inizio.

Per definire una nuova variabile per le date in ANALYSIS si procede come segue:

DEFINE <nome variabile> <tipo di data>

il tipo di data deve essere scelto fra quelli presentati sopra, ad esclusione però di quelli contenenti "today". Con le variabili di date è possibile compiere operazioni aritmetiche, il risultato delle quali sarà espresso in numero di giorni. Supponiamo, per esempio, di avere le variabili DATAINSORG e GUARIGIONE nel file corrente. Possiamo definire una nuova variabile, per esempio DURATA, cui verranno assegnati i valori risultanti dalla differenza fra la data di guarigione e la data dell'insorgenza della malattia. Ecco come procedere:

DEFINE DURATA

LET DURATA = GUARIGIONE - DATAINSORG

In questo modo la variabile DURATA conterrà i numeri di giorni della durata della malattia. Nel file le date sono trattate come variabili di testo pertanto, se si compongono comandi contenenti direttamente date, è necessario racchiuderle fra virgolette:

LET DURATA = "31/12/91" - DATAINSORG

Negli esempi precedenti le durate sono espresse in giorni. È possibile però esprimerle in settimane, o in anni, semplicemente dividendole per 7 e 365,25, rispettivamente:

DEFINE SETTIMANE

DEFINE ANNI

LET SETTIMANE = DURATA/7 ["/" indica la divisione]

LET ANNI = DURATA/365,25

Esprimendo le durate in termini di settimane è desiderabile che queste siano arrotondate per evitare la presenza di decimali. Si usa allora, al posto del segno "/" della divisione, l'operatore DIV, il quale eliminerà i giorni in più che dovessero risultare da un resto della divisione diverso da zero:

LET SETTIMANE = DURATA DIV 7

Altre informazioni per ulteriori comandi eseguibili con le date sono forniti nel Manuale elettronico, capitolo 36, nella sezione dedicata al comando LET.

LABORATORIO: SCRIVIAMO UN PROGRAMMA

Carichiamo ancora il file ORIENTE.REC, contenente i dati di una tossinfezione alimentare, in ANALYSIS, con il comando READ ORIENTE, ed eseguiamo alcune operazioni descritte in questo capitolo.

```
SELECT MALATO = "Y"
```

Con questo comando abbiamo selezionato soltanto i malati. Infatti se eseguiamo il comando `FREQ MALATO` avremo solo gli individui con il valore "+".

Bisogna usare sempre le virgolette se selezioniamo condizioni in campi carattere (non numeriche) e se usiamo lettere maiuscole o minuscole. Non servono invece le virgolette nei campi numerici.

Per selezionare più condizioni è sufficiente eseguire più comandi successivi.

```
SELECT MALATO = "Y"
```

```
SELECT ETA < 12
```

il programma in questo caso concatena i due comandi aggiungendo "AND" tra le due condizioni includendole tra parentesi, come si può vedere in alto sullo schermo, in corrispondenza di "Criteria" (Nota: con l'operatore AND sia la prima che la seconda condizione devono essere vere). Quindi i comandi SELECT eseguiti in successione si sommano come se fosse stato introdotto l'operatore logico AND. Naturalmente si possono scrivere più condizioni su una sola riga utilizzando AND o altri operatori logici.

Le condizioni di selezione possono essere cancellate sia caricando un altro file (READ), sia lanciando il comando SELECT senza alcuna condizione.

```
SELECT
```

```
SELECT (ETA > 1) AND (ETA < 15)
```

```
FREQ ETA
```

L'operatore logico AND - *coniunzione* - è vera quando sono vere *tutte* le condizioni che sono congiunte; il comando selezionerà allora tutti i record per i quali sono veri l'età maggiore di 1 e minore di 12 anni .

Le parentesi sono necessarie per separare le condizioni ed è più sicuro usarle sempre per evitare errori. Fate attenzione inoltre alla logica: i seguenti comandi non selezioneranno alcun record da analizzare.

```
SELECT SINTOMI = "Y"
```

```
SELECT SINTOMI = "N"
```

Infatti non c'è nessun record con il valore "Y" per la variabile SINTOMI e *contemporaneamente* con il valore "N" nello stesso campo.

```
SELECT
```

```
SELECT (SINTOMI = "Y") OR (SINTOMI = "N")
```

```
FREQ SINTOMI
```


Questo comando selezionerà *sia* i record con sintomi, *sia* quelli senza. L'operatore logico OR - *disgiunzione* - rende vera una proposizione quando *almeno una* delle condizioni contenute è vera.

SELECT può anche essere usato per escludere record. Se si vogliono escludere i record con valore "Y" della variabile RICOVERO, per esempio per calcolare la durata media dei sintomi tra i non ricoverati, è sufficiente eseguire questo comando:

SELECT

SELECT RICOVERO < > "Y"

FREQ DURATASINT

i segni "< >" significano o minore o maggiore, dunque: "non uguale a"

Passiamo ora al comando ROUTE , per stampare i risultati oppure trasferirli in un file. Digitando F5 attiviamo la stampante; possiamo notare come nella barra di stato la scritta "Printer on" si trasforma in "Printer off" e sullo schermo appare "output - printer" al posto di "output - screen". Tutto ciò indica che il prodotto delle successive elaborazioni verrà trasferito sia allo schermo che alla stampante. Inoltre, sempre dopo aver digitato F5, sullo schermo apparirà anche ROUTE PRINTER, questo è il comando che corrisponde a <F5>.

Il comando ROUTE può essere utilizzato anche per trasferire i risultati del nostro lavoro in un file, in modo da poterlo rivedere e correggere in EPED.

ROUTE INDAGINE.DOC

con questo comando Analysis trascriverà sul file Indagine.doc il prodotto delle elaborazioni che si eseguiranno da qui in avanti, tranne i grafici.

SORT ETA

ordina i record in ordine di età, dalla minore alla maggiore, usando la variabile ETÀ. Eseguito poi:

LIST MALATO SESSO ETÀ RICOVERO

si vedrà che il numero dei record non sarà più in ordine di introduzione, ma lo saranno in base all'età degli ospiti.

Una volta compiuto un lavoro di elaborazione si possono salvare, in un file, tutti i comandi che si sono eseguiti fino ad un limite di 20 righe. Usando il comando SAVE, tutti i comandi eseguiti verranno salvati in un file che dovrà avere estensione .PGM .

SAVE INDAGINE.PGM

Il file così registrato potrà essere letto, rivisto, modificato e corretto in EPED. Provate a visualizzare INDAGINE.PGM in EPED, aggiungendo magari qualche comando. Tornate poi in ANALYSIS e digitate:

RUN INDAGINE

Potrete vedere come il programma esegue in pochi secondi , ed esattamente, tutti i comandi di questa esercitazione.

Per ulteriori dettagli si vedano i capitoli 11, 13 e 36 del manuale elettronico.

Capitolo 10

ANCORA SU ANALYSIS

MODIFICARE ALCUNE IMPOSTAZIONI DI ANALYSIS

In ANALYSIS ci sono alcuni comandi utili al controllo e al perfezionamento della forma finale dei lavori su Epi Info. Tali comandi permettono, per esempio, di inserire titoli, commenti e valori delle variabili; di modificare la grandezza della pagina, la posizione delle righe e permettono di produrre nuovi file nei formati specificati. La maggior parte di questi comandi non è necessaria per usare Epi Info nelle indagini epidemiologiche; sono state però incluse per essere usate nei sistemi permanenti che richiedono opzioni particolari e per i quali vale la pena dedicare più tempo alla programmazione.

I COMANDI SET PER LA STAMPANTE E I GRAFICI

Il comando SET serve a controllare numerose impostazioni che influenzano i risultati prodotti da Epi Info. Permette di inserire titoli, commenti e valori di variabili nella relazione finale; di controllare la dimensione della pagina ed il posizionamento delle righe e di produrre nuovi file in formati specificati. La maggior parte di tali comandi non sono necessari per usare Epi Info nelle indagini epidemiologiche; sono inclusi per l'uso in sistemi permanenti in cui sono richieste particolari caratteristiche ed una spesa di tempo extra per "programmare", se ne vale la pena. Le uniche impostazioni importanti (SET=Imposta) sono quelle per la stampante e i grafici; per tutte le altre si rimanda al capitolo 36 del Manuale elettronico, alla voce SET.

I grafici in ANALYSIS necessitano di "driver" per essere stampati su stampanti diverse. I driver (file .BGI) nella Versione 6 di Epi Info sono diversi da quelli usati con la Versione 5, e hanno nomi che iniziano con "\$", ad esempio, \$OKI92 per le stampanti Okidata 92. Per impostare una stampante si deve eseguire il comando seguente:

SET PRINTER = <nome driver stampante>

il nome del driver della stampante va scelto fra i driver elencati di seguito:

Nome	Descrizione	Nome	Descrizione
BMP	BMP (Windows 3.x) Format	HP7550	HP Plotter 7550
CGM	Computer Graphics Metafile	HP7585	HP Plotter 7585
FX	Epson FX80 Printer	HP7595	HP Plotter 7595
PS	Postscript	IBMQ	IBM Quietwriter Printer
TIF	Compressed TIF Format	IMG	GEM IMG Format
UTIF	Uncompressed TIF Format	LJ	Hewlett-Packard Laserjet
WPG	WordPerfect Graphics	LJ3R	Hewlett-Packard Laserjet III
CFX	Epson 9-pin Color Printer	LQ	Epson 24-Pin Printer
CLQ	Epson 24-pin Color Printer	OKI92	Okidata 92 Printer
DJ	Hewlett-Packard Deskjet	PCX	PCX (640 x 480) Format
DJ	Hewlett-Packard Color Deskjet	PJ	Hewlett Packard Paintjet
DXF	AutoCad DXF Format	PP24	IBM Proprinter, 24 Pin

HP7090	Hewlett-Packard Plotter 7090	TSH	Toshiba 24-Pin Printer
HP7470	HP Plotter 7470	VSHO	ANSI NAPLPS/Video Show Format
HP7475	HP Plotter 7475		

per impostare una stampante HP Laserjet è quindi sufficiente eseguire il comando:

SET PRINTER = LJ

Per stampare i grafici su file è necessario impostare un driver apposito. Si tratta dei primi due driver della tabella mostrata sopra (CGM e BMP). Impostando questi driver, ANALYSIS produrrà un file .BMP (Bitmap) o .CGM (Computer Graphics Metafile) che potranno essere utilizzati direttamente in tutti i programmi per WINDOWS. Per impostare la stampa dei grafici su file digitare:

SET GRAPH = BMP oppure SET GRAPH = CGM

Dopo aver impostato la stampante o i grafici, digitare **SET: ANALYSIS** mostrerà tutte le impostazioni correnti. In fondo ci sarà un riga con "printer type:" (tipo stampante) e un'altra con "GRAPHtofile =" (tipo file grafico). Controllare se compaiono le impostazioni che si sono inserite.

CREARE UN NUOVO FILE EPI INFO

A volte è utile creare un nuovo file Epi Info contenente record selezionati o modificati in seguito a procedure particolari, oppure contenenti variabili definite e altri cambiamenti temporanei. Questo lavoro può essere compiuto mediante i comandi ROUTE e WRITE RECFILE. Per esempio, per creare un file chiamato SELEZION.REC a partire dai dati del file di diverso nome che si è modificato, si usa il seguente comando:

ROUTE SELEZION.REC

WRITE RECFILE

Il comando Route manda i dati in uscita in un nuovo file dal nome prescelto, ed è importante far seguire questo nome dal suffisso .REC per poter essere riconosciuto da Epi Info. Il comando WRITE seguito da RECFILE serve per creare un nuovo file di formato .REC di Epi Info. Questo comando può essere seguito da una lista di nomi di campi, separati da spazi, in modo da includere nel nuovo file soltanto questi campi. Nel nuovo file saranno inoltre inseriti soltanto i record attualmente attivi (selezionati con SELECT).

I record cancellati, quelli cioè cancellati provvisoriamente nel programma ENTER, saranno omessi nel nuovo file se è attivo il comando: SET PROCESS = UNDELETED (processa solo i non cancellati). Se invece è stato dato il comando: SET PROCESS = DELETED, saranno scritti soltanto i record cancellati. Per inviare nel nuovo file tutti i record, indipendentemente dal fatto se siano stati cancellati o meno, si usa il comando: SET PROCESS BOTH (processa entrambi). Normalmente ANALYSIS processa soltanto i record non cancellati.

Il comando ROUTE SCREEN pone termine alla scrittura del file e i dati elaborati vengono nuovamente indirizzati al solo schermo. La scrittura dei file si interrompe anche premendo <F10> per uscire da ANALYSIS.

RISTRUTTURARE I RECORD

Il comando WRITE può essere usato per cambiare il numero o l'ordine delle variabili di ciascun record, per scrivere in un file le variabili definite, oppure per accorciare o allungare una variabile. Come descritto sopra, il primo passo da compiere è indirizzare (ROUTE) i dati al file appropriato. Se il file deve essere del formato di Epi Info il comando write deve essere seguito da RECFILE, altrimenti sarà creato un file di testo normale (formato ASCII). Si tenga presente che se si desidera creare proprio un normale file di testo, questo non sarà chiuso automaticamente da ANALYSIS, ma sarà necessario dare il comando ROUTE SCREEN per avvertire il programma che la scrittura del file è terminata e che deve chiudere il file.

Supponiamo che il file attualmente attivo in ANALYSIS contenga i campi MALATTIA, ETA, CITTA e SESSO; e supponiamo di voler creare un nuovo file contenente questi stessi campi in ordine differente, e di voler aggiungere un nuovo campo di dieci caratteri, riempito con dei 9, che indicherà i campi i cui dati sono mancanti (blank entries). I comandi da eseguire per avere questi risultati sono:

```
DEFINE MANCANTI _____
```

```
MANCANTI = "9999999999"
```

```
ERASE NUOVO.REC [Cancella NUOVO.REC, nel caso esista già]
```

```
ROUTE NUOVO.REC [Crea il file NUOVO.REC]
```

```
WRITE RECFILE Citta:2 Sesso:1 Malattia:3 Eta:2 Mancanti
```

```
ROUTE SCREEN [Chiusura del file. Comando facoltativo]
```

Grazie al comando RECFILE il file sarà del formato Epi Info. Si noti che gli elementi che seguono questo comando (città, sesso, ecc.) sono separati da spazi e i numeri posti a fianco di ciascuno di essi, dopo i due punti, indicano la larghezza del campo. Si noti inoltre che "9999999999" deve essere scritto nel file .REC come nome di variabile.

Supponiamo adesso di aver bisogno di un file ASCII particolare, con i campi delimitati da spazi, che è il formato richiesto dai programmi di regressione logistica [ad es. in S.A.S. o S.P.S.S]. Il file va scritto utilizzando i seguenti comandi:

```
ERASE REGRESS.DAT [Cancella REGRESS.DAT, nel caso esista già]
```

```
ROUTE REGRESS.DAT [Crea il file REGRESS.DAT]
```

```
WRITE MALATTIA " " ETA " " SESSO " " CITTA " "
```

```
ROUTE SCREEN [Comando di chiusura del file. Obbligatorio]
```

Nella sezione dedicata al comando WRITE nel Manuale elettronico di Epi Info (cap. 36), si troverà anche il metodo per specificare il numero e la collocazione dei decimali e il metodo per inserire caratteri speciali come <Invio>, utili per produrre etichette e indirizzi per lettere.

I COMANDI ??, TYPE, ECHO, BREACK, IMMEDIATE

Scrivendo programmi che devono essere usati da altre persone, molto spesso è necessario inserire domande da indirizzare all'operatore. Per esempio, se si scrive un programma chiamato SETTIMANA con lo scopo di elaborare i dati settimanali di un sistema di sorveglianza, si avrà la necessità di chiedere all'operatore per quale settimana si desidera che siano elaborati i dati, oppure per quale periodo compreso fra due date. Per fare questo si inserisce un messaggio per l'operatore racchiuso fra due punti interrogativi:

SETTIMANA = ?Quale settimana si vuole analizzare (1-52):?

Questo messaggio apparirà sullo schermo e la risposta dell'operatore sarà introdotta nel campo della variabile SETTIMANA. Questo stesso meccanismo può essere utilizzato anche per altri comandi che necessitano di valori variabili. Per esempio:

LIST ?Di quali campi si desidera l'elenco (nomi separati da spazi):?

TABLES ILL ?Nome della variabile di esposizione:?

Il comando TYPE può essere usato per far apparire messaggi sullo schermo, nel corso della stampa o in un file, a seconda del comando ROUTE dato. In questi messaggi possono essere inseriti anche valori variabili utilizzando il segno "@", come in:

TYPE "\C Rapporto settimanale della settimana @SETTIMANA"

dove @SETTIMANA prende il valore del numero della settimana prescelta, e il segno "\C" significa che il messaggio apparirà al centro dello schermo. Per esempio se la settimana scelta è la numero 14 il messaggio apparirà come segue:

Rapporto settimanale della settimana 14

Il comando TYPE ha anche un'altra funzione. **TYPE TESTO.TXT**, per esempio, invierà il contenuto del file TESTO.TXT allo schermo o alla stampante, a seconda del comando ROUTE dato. Questo comando è utile per esaminare il contenuto di un file di programma o per aggiungere dei testi presi da file preesistenti in una relazione che si sta scrivendo.

Il comando ECHO è simile all'omonimo comando del DOS: mostra solo sullo schermo il testo che segue il comando. Per esempio **ECHO ARRIVEDERCI**, farà apparire sullo schermo la parola ARRIVEDERCI; scrivendo però **TYPE "ARRIVEDERCI"**, la parola sarà indirizzata allo schermo, alla stampante o a un file, a seconda del comando ROUTE imposto. ECHO è utile quando si scrivono programmi e si vogliono inserire istruzioni o messaggi per l'operatore che compaiano solo sullo schermo e non nei file o negli stampati.

Ai comandi TYPE e TITLE si possono aggiungere diversi segni speciali per ottenere determinati effetti particolari. Uno di questi è il segno "\n" che significa "nuova linea" e serve per mandare a capo il testo nel punto dove è inserito. Per esempio:

TYPE "Rapporto mensile \n Giugno 1992"

farà comparire il messaggio sullo schermo su due righe:

Rapporto mensile

Giugno 1992

Il comando BREAK esegue un comando predefinito ogni volta che cambia il valore di un campo particolare, o di un gruppo di campi, durante lo scorrimento di un elenco (comando LIST). Normalmente il comando predefinito è il cambiamento automatico

della pagina, in modo da elencare nella stessa pagina ogni gruppo di campi con una stessa caratteristica (per esempio la stessa città o la stessa età). Per ottenere questo risultato bisogna innanzitutto ordinare il file raggruppando i record con una caratteristica identica (comando SORT). Si dà quindi il comando **BREAK CITTA** (o **BREAK ETA**, ecc.) e infine si elencano i record (comando LIST). Ulteriori dettagli per impostare il comando BREAK in modo che esegua qualsiasi altra operazione di Epi Info oltre al cambiamento di pagina, si trovano nel capitolo 36 del Manuale elettronico, sotto le voci BREAK e SET.

Il comando **IMMEDIATE** deve essere seguito da un altro comando di quelli che sono eseguiti per ciascun record contenuto nel file. **IMMEDIATE** ha l'effetto di far eseguire il comando che segue una sola volta invece che ripetutamente per tutti i record. Questo è particolarmente utile per i comandi IF e LET. Ecco due esempi:

```
IMMEDIATE IF QUARTIERE = 5 THEN TYPE CINQUE.TXT
```

```
IMMEDIATE ANNO = ?Anno da analizzare?
```

Ulteriori dettagli sulle possibilità di interazione con lo schermo sono contenute nel capitolo 36 del Manuale elettronico di Epi Info.

RECORD RIASSUNTIVI: OUTPUT, SUMFREQ E SUMTABLES

In molte situazioni è richiesto l'uso di record riassuntivi che racchiudano i risultati di più persone o casi. È comodo inoltre essere in grado di usare i risultati di una tabella Epi Info e condurre ulteriori indagini a partire dalla tabella stessa. Un'Unità Sanitaria Locale potrebbe, per esempio, elaborare un rapporto mensile sulle malattie individuali e scrivere i risultati in un file riassuntivo. Questo file potrebbe poi essere spedito alle autorità sanitarie comunali, provinciali o regionali, le quali inserirebbero in un unico file riassuntivo tutti i risultati parziali provenienti dalle periferie, in modo da possedere un quadro riassuntivo delle malattie individuali a livello comunale, provinciale o regionale.

Il comando OUTPUT, usato insieme con i comandi TABLES o FREQ, produce un record di dati di formato Epi Info a partire da ciascuna casella di una tabella. Combinando questi comandi con il comando ROUTE si crea un file Epi Info riassuntivo che può essere indirizzato alla stampante o scritto sul disco. Ecco un esempio di comando:

```
ROUTE SOMMARIO.REC
```

```
OUTPUT TABLES MALATTIA CITTA
```

Questo comando crea un file dal nome SOMMARIO.REC contenente record con la somma dei casi di malattia per ciascuna città. Il file avrà lo standard Epi Info (.REC) e avrà le seguenti tre variabili:

```
MALATTIA CITTA COUNT
```

La variabile COUNT (conteggio) è aggiunta dal programma e per ciascun record questa variabile assume il valore di una casella della tabella prodotta dal comando TABLES MALATTIA CITTA. Per esempio un record potrebbe contenere i valori: "Epatite", "Ravenna" e 12; ciò significa che nella suddetta tabella sono stati contati 12 casi nella casella corrispondente ai casi di epatite a Ravenna.

Il risultato dell'esempio appena illustrato sarebbe molto più utile se fosse indicato anche il numero della settimana cui si riferiscono le insorgenze della malattia. Questo può essere fatto aggiungendo la variabile NUMSETT come nell'esempio che segue:

OUTPUT TABLES MALATTIA CITTA NUMSETT

I comandi SUMFREQ e SUMTABLES sono usati per elaborare ulteriormente i record prodotti dal comando OUTPUT. Questi due comandi producono tabelle simili a quelle prodotte da TABLES e FREQ, ma le loro caselle presentano risultati riassuntivi al posto dei conteggi dei record.

I comandi FREQ e TABLES *contano* il numero di record che soddisfano determinate condizioni e mettono il risultato nell'apposita casella. SUMFREQ e SUMTABLES *sommano* i valori della prima variabile specificata nel comando corrispondente, alle altre variabili e posizionano il risultato nelle caselle della tavola.

I comandi OUTPUT e SUMTABLES operati di seguito costituiscono un'ottima procedura per ottenere e archiviare i file riassuntivi, in modo che non ci sia bisogno di ripetere da capo le analisi di un file, specie se di ampie dimensioni, ogni volta che si desidera operare ulteriori sintesi dei dati. Questi comandi sono inoltre essenziali per i sistemi di sorveglianza che utilizzano solo dati riassuntivi, per esempio gli osservatori epidemiologici a livello regionale o statale. Nel capitolo 26 del Manuale elettronico di Epi Info si troveranno molti esempi sull'uso di questi comandi, nell'ambito di un sistema di sorveglianza che utilizza dati riassuntivi e aggregati.

GRAFICI DI RECORD RIASSUNTIVI

A volte è utile disegnare un grafico di dati riassuntivi in cui sia rappresentata la somma dei valori di ogni record, piuttosto che il conteggio dei record riportanti lo stesso valore. I record dei dati riassuntivi possono essere prodotti in due modi:

1. Introducendo i dati riassuntivi direttamente nel questionario. Una variabile potrebbe essere MALATTIA, mentre i dati riassuntivi dell'insorgenza di malattie in diverse città, o in un certo periodo di tempo, potrebbero essere inseriti nella variabile CASI. I record si presenterebbero allora come:

<u>MALATTIA</u>	<u>CASI</u>
EPATITE A	600
MALARIA	412
EPATITE A	314

2. Usando il comando OUTPUT TABLES che produce record riassuntivi. Un riassunto dei dati individuali contenuti in una base di dati potrebbe essere prodotto con il comando OUTPUT TABLES MALATTIA.

Per produrre un grafico con la somma dei valori nella variabile B (ordinata) e con le varie malattie nell'ascissa (variabile A), si usa la seguente sintassi dei comandi:

BAR <Variabile A>/SUM = <Variabile B>

Nell'esempio sopra, il comando corretto per produrre un grafico a barre è:

BAR MALATTIA/SUM = CASI

Le due barre del grafico mostreranno 914 casi di EPATITE A e 412 casi di MALARIA.

UNIRE I DATI IN CAMPI MULTIPLI NELLO STESSO FILE: IL COMANDO COMBINE

COMBINE consente che i vari campi vengano trattati come se fossero un campo unico. Ipotizzate, ad esempio, che abbiate vari campi in un questionario basato su casi clinici come segue:

Diagnosi1 _____

Diagnosi2 _____

Diagnosi3 _____

Diagnosi4 _____

Vi interessa sapere quante volte "polmonite" appare come diagnosi tra tutti i casi. Visto che potrebbe trovarsi in qualsiasi dei quattro campi, sarebbe necessario scrivere un'istruzione IF per contare la parola 'polmonite' se essa appare in uno qualsiasi dei campi, oppure addirittura un programma per creare dei record nuovi per ogni diagnosi piuttosto che per ciascun "caso." Tecnicamente questa si chiama *normalizzazione* del database. Si potrebbe definirla come lo spostamento di tutti i dati che si desidera contare o sommare nello stesso campo.

COMBINE risolve il problema dei dati che vengono automaticamente distribuiti in campi multipli. Quando inserite i comandi

DEFINE DIAGNOSI _____

COMBINE VALUE = DIAGNOSI1 TO DIAGNOSI4

i comandi successivi **FREQ** e **TABLES** che comprendono **DIAGNOSI** simuleranno un file con quattro volte il numero di record contenuti nel file originale. Ci sarà un record caso per ciascuno dei quattro campi di diagnosi. Nel primo, **DIAGNOSI** conterrà i dati per **DIAGNOSI1**, nel secondo per **DIAGNOSI2**, ecc. Abbiamo ora un file simulato basato sul numero di diagnosi (comprese quelle non compilate). Il comando **FREQ DIAGNOSI** può anche produrre un conteggio del tipo:

Amebiasi	5
Polmonite	12
Epatite	1

Un comando come **TABLES ETA DIAGNOSI** produrrà una tabella con il seguente formato, dove **XX** rappresenta una cella individuale:

		DIAGNOSI			
		XX	XX	XX	XX
ETA		XX	XX	XX	XX
		XX	XX	XX	XX
		XX	XX	XX	XX
					TOTAL

Sono stati uniti (COMBINE) i campi **DIAGNOSI1..DIAGNOSI4**. Il numero effettivo di record potrebbe essere diverso da quello che risulta dalla tabella o dal conteggio, dato che ciascun dato immesso, da **DIAGNOSI1** a **DIAGNOSI4**, aggiunge un record alla tabella. Se **SET IGNOREMISSING** è su **OFF**, allora il numero di "record" sarà esattamente

quattro volte il numero dei record effettivi; altrimenti, i valori mancanti non verranno aggiunti.

TOTAL rappresenta il numero di diagnosi quando SET IGNOREMISSING è su ON, e quattro volte il numero di record (perché ci sono quattro campi diagnostici) quando IGNOREMISSING è su OFF.

È possibile attivare soltanto un comando COMBINE per volta. L'inserimento di un secondo comando COMBINE, disinscrive i COMBINE precedenti. La parola COMBINE da sola disattiva la funzione COMBINE.

Dopo aver dato il comando COMBINE, WRITE RECFILE produrrà un file con un record per ciascuna diagnosi (IGNOREMISSING ON) o per ogni campo contenuto nell'istruzione COMBINE (IGNOREMISSING OFF). Nel nostro esempio, con IGNOREMISSING OFF, verranno prodotti quattro volte il numero di record contenuti nel file originale. Ciascuno avrà una copia di tutti gli altri campi (come NOME, SESSO, INDIRIZZO) contenuti nei record corrispondenti nel file dati originale.

Talvolta sarà importante includere il nome dei campi che sono stati uniti nei nuovi record in modo che si possa costruire una tabella. Nel caso di una tossinfezione alimentare, potrebbero apparire i seguenti campi:

```

Pollo      <Y>
Manzo     <Y>
Insalata  <Y>

```

Se interessa fare un grafico o un conteggio delle esposizioni per ogni tipo di cibo, invece di tre campi separati conviene costruirne un solo campo, di nome CIBO, contenente i nomi dei cibi ingeriti. Invece dei record:

NOME	POLLO	MANZO	INSALATA
Jones	Y	N	Y
Smith	N	Y	N

Si avranno i record:

NOME	CIBO	INGERITO
Jones	Pollo	Y
Jones	Manzo	N
Jones	Insalata	Y
Smith	Pollo	N
Smith	Manzo	Y
Smith	Insalata	Y

in modo che il comando:

TABLES CIBO INGERITO

produrrà la seguente tabella:

	INGERITO		
CIBO	+	-	
Pollo	1	1	2
Manzo	1	1	2
Insalata	2	0	2
	4	2	6

Con SET PERCENTS = ON, questa tabella darà i tassi di attacco per i vari cibi. Se interessa costruire un grafico del numero di esposizioni ai vari cibi, si potrebbe usare:

```
SELECT INGERITO = "Y"
```

BAR CIBO

per collocare i nomi dei cibi lungo l'asse orizzontale e il numero di persone sull'asse verticale di un grafico a barre. Ipotizziamo un questionario con i seguenti campi:

Numero di casi di MALARIA ##### DIARREA ##### NEFRITE #####.

È possibile combinare i campi nel seguente modo:

```
DEFINE CONTA #####
```

```
DEFINE MALATTIA _____
```

```
COMBINE NAME=MALATTIA VALUE=COUNT MALARIA DIARREA  
NEFRITE
```

Se si desidera un codice invece dei nomi dei campi MALARIA, DIARREA e NEFRITE, si usa la stessa sintassi dell'istruzione RECODE per la ricodifica dei campi:

```
COMBINE NAME=MALATTIA VALUE=COUNT MALARIA=1 DIARREA=2  
NEFRITE=3
```

Per produrre la tabella finale, con i conteggi delle malattie, si userà il seguente comando

```
SUMTABLES COUNT DISEASE
```

Per la sintassi esatta di COMBINE vedere il capitolo 36 del Manuale elettronico alla voce corrispondente.

CONTROLLO DEL PROGRAMMA: I COMANDI RUN, GOTO E RETURN

Questi tre comandi vanno inseriti nella scrittura dei programmi per ottenere alcuni effetti particolari, fra i quali quelli descritti di seguito. Per notizie più dettagliate si rimanda come al solito al capitolo 36 del Manuale elettronico.

Quando è stato completato un certo programma può essere utile in certi casi avviare automaticamente un altro programma. Questo lavoro può essere compiuto utilizzando il comando RUN (AVVIA).

Può anche essere utile saltare un pezzo di un certo programma in determinate circostanze (se alcune condizioni sono vere). Lo si può fare usando il comando GOTO (VAI A), il quale va direttamente alla riga specificata all'interno dello stesso programma.

Una sequenza di righe del programma può dover essere eseguita più di una volta. Se la sequenza è eseguita più volte nello stesso punto del programma è chiamata "loop" (circolo), perché il programma ricomincia dalla prima riga della sequenza dopo aver eseguito l'ultima. I "loop" sono descritti più in dettaglio nel capitolo 36 del Manuale elettronico, sotto la voce GOTO. Le condizioni per i "loop" sono poste dal comando IMMEDIATE IF.

Se una stessa sequenza di righe del programma è richiamata varie volte in punti diversi del programma, è più utile inserire questa sequenza in una subroutine da richiamare con il comando GOTO. Il comando RETURN indica la fine della subroutine e provoca il

ritorno alla riga che segue immediatamente quella contenente il comando GOTO. Le subroutine sono descritte alla voce RETURN nel capitolo 36 del Manuale elettronico.

LABORATORIO: IL PROGRAMMA ORIENTE.PGM

Nel quarto dischetto allegato al presente manuale trovate un programma riferito alla Tossinfezione alimentare all'Hotel Oriente. Avviamo il programma in ANALYSIS digitando il comando RUN al prompt di Epi Info: EPI>. Dopo aver digitato INVIO osservate i grafici e le tabelle che compaiono sullo schermo.

Alla fine, se avviamo EPED: possiamo aprire (<F2> Open file in this window) il file ORIENTE.PGM.

Rappresenta un compilato dei comandi che avete visto poco fa in sequenza.

Analizzate i comandi e metteteli in relazione con i risultati dell'elaborazione. Osservate la sintassi dei comandi e la loro sequenza. Potete provare ad apportare delle modifiche ai comandi e alla loro sequenza per osservare, in ANALYSIS con il comando RUN, i risultati delle vostre variazioni. Questo è il sistema usato da tutti i programmatori del mondo per imparare a governare il programma.

Per ulteriori dettagli si vedano i capitoli 11, 13 e 36 del Manuale elettronico.

Capitolo 11

STATCALC ED EPITABLE: CALCOLATORI STATISTICI

INTRODUZIONE

STATCALC realizza statistiche per dati inseriti dalla tastiera in tabelle mostrate sullo schermo. Vengono fornite facilitazioni per realizzare analisi di tabelle 2×2 singole e stratificate, per calcolare le dimensioni del campione per studi comportanti proporzioni e per analizzare andamenti (trends) con o senza stratificazione.

AVVIO DI STATCALC

Avviare STATCALC dal menù principale di EPI6. Apparirà un menù contenente le seguenti scelte:

- Tables (2×2 , $2 \times n$)
- Sample size & power
- Chi square for trend

Per scegliere, muovete la barra sull'opzione con i tasti freccia e premete <Invio>, oppure premere la lettera evidenziata, ad esempio T per la prima opzione. La pressione di <F1> mostra schermi di ausilio contenenti informazioni circa il programma ed i metodi statistici. Vi raccomandiamo di leggerli attentamente.

ANALISI DI TABELLE SINGOLE E STRATIFICATE (2×2 fino a 2×10)

Questa opzione presenta una tabella 2×2 sullo schermo. Per realizzare statistiche per una singola tabella 2×2 , si immettono semplicemente quattro numeri nella tabella e quindi si preme <F4> per realizzare i calcoli. STATCALC calcola l'odds ratio ed il rischio relativo con i limiti di confidenza, tre tipi di test del chi-quadrato e, se opportuno, il test esatto di Fisher. Dopo che questi sono apparsi, è ancora possibile scegliere di calcolare gli esatti limiti di confidenza. Ciò è molto utile se i numeri nella tabella sono piccoli.

Per tabelle 2×2 , possono essere inseriti strati addizionali premendo <F2> dopo che sono comparse le statistiche. Dopo che sono state introdotte tutte le stratificazioni, la pressione di <F4> mostrerà i risultati dell'analisi stratificata e la possibilità di calcolare gli esatti limiti di confidenza, se lo si desidera.

Per tabelle singole che hanno fino a un massimo di 10 caselle sulla verticale, continuare semplicemente ad inserire numeri, seguiti da un ulteriore <Invio> di conferma. STATCALC calcolerà un chi-quadrato per la tabella $2 \times n$.

Le statistiche prodotte con STATCALC sono simili a quelle realizzate dai file di dati col programma ANALYSIS, con l'eccezione che STATCALC calcola sia i limiti di confidenza di Cornfield sia gli esatti limiti di confidenza per tabelle singole e stratificate. I limiti di confidenza per il rischio relativo nelle tabelle singole e stratificate sono quelli di Gre-

enland e Robins (Biometrics 1985; 41:55-68). Vengono forniti gli esatti limiti di confidenza basati sull'algoritmo di Mehta, Patel e Gray (J Am Statistical Assoc 1985;80:969-973) da utilizzare con le tabelle in cui i limiti di confidenza di Cornfield non sono sufficientemente precisi (tabelle di contingenza con pochi dati). Nel capitolo 32 del manuale elettronico vengono fornite ulteriori informazioni sui metodi statistici.

DIMENSIONE DEL CAMPIONE E POTENZA

Selezionando "Sample size & power," viene presentato un menù secondario con le seguenti scelte:

Population survey

Cohort or cross-sectional

Unmatched case-control

Tutti i calcoli sono destinati a studi in cui i risultati sono proporzioni espresse in percentuale. In uno studio di sorveglianza della popolazione o descrittivo, trovare la proporzione delle persone che rispondono SI ad una determinata domanda è un tipico obiettivo. Il programma chiede di specificare le probabilità che il proprio campione rappresenti la situazione reale nella o nelle popolazioni sotto esame e le imprecisioni che si è disposti a tollerare; ed infine calcolerà la dimensione del campione basandosi su questi presupposti: campionamento casuale della popolazione, 100% di partecipazione, 100% di accuratezza nell'ottenere e registrare le risposte ed altre condizioni che sono raramente presenti nella vita reale. Se ci si attende il 90% di partecipazione piuttosto che il 100%, si dovrà moltiplicare la dimensione del campione per 100/90 per ottenere una stima più realistica.

Le dimensioni del campione calcolate sono solo indicative e, in molti studi, il numero dei casi a disposizione, o delle interviste che possono essere realizzate restando nei limiti di spesa previsti, sono i fattori determinanti della decisioni. Questi programmi forniscono un modo utile per ottenere stime sulla base delle quali possono essere pesati altri fattori. Nel corso dello studio gli assunti circa i risultati ipotetici saranno rimpiazzati dai risultati attuali ed i livelli di confidenza accettati saranno rimpiazzati dai valori di p o dai limiti di confidenza derivati dai risultati. Fino a che questo senso del poi non è disponibile, i calcoli della dimensione del campione permettono, in parte, una previsione.

L'operazione relativa a questi calcoli procede mediante l'individuazione del livello di confidenza, della potenza e degli assunti relativi alla stima della proporzione attuale, del rischio relativo o dell'odds ratio più vicini a quelli desiderati. Per ultimo, si sceglie il valore più lontano dal valore della popolazione reale che si è disposti ad accettare. Per esempio, nello studio Prospettivo o Cross-Sectional, supponiamo di scegliere il 10% come proporzione ipotetica di risposte positive ad una domanda posta alle persone non esposte. Si potrebbe voler inserire il 20% come valore più vicino che può essere individuato nelle persone esposte, intendendo che qualsiasi valore uguale o maggiore del 20% darebbe un valore p , nello studio finale, uguale o inferiore a 0,05. Ciò corrisponde ad un rischio relativo di 2,0 o ad una odds ratio di 0,5. Ognuno di questi tre valori può essere immesso inizialmente per ottenere gli stessi risultati.

Dopo che gli assunti sono stati introdotti, premendo <F4> verranno effettuati i calcoli. Il risultato per l'assunto esatto introdotto sarà evidenziato e verrà inoltre presentata una serie di altri valori basati sulla variazione degli assunti. I risultati possono essere stampati o inviati ad un file, come descritto più avanti.

CHI-QUADRATO PER TREND

Il test chi quadrato per trend viene spesso utilizzato per studi dose-risposta e può anche essere usato per valutare andamenti per età, tempo o qualunque variabile ordinata. I dati possono essere stratificati su altre variabili come età e sesso per eliminare confondimenti derivanti da queste variabili.

Il chi-quadrato esteso di Mantel-Haenszel che viene calcolato, riflette la deviazione di un andamento lineare dall'orizzontale (cioè nessun trend). Se il valore di p associato è minore di 0,05, vi è il 95% di probabilità che un andamento sia presente nella popolazione presa in considerazione.

I dati da analizzare mediante questo test debbono essere raggruppati secondo una sequenza numerica. I gruppi più semplici sono arbitrariamente quelli di 0, 1, 2, 3, ecc. Queste sono le "osservazioni" o i "punteggi". Altri punteggi possono essere i punti medi dei gruppi, come 0 per 0 bicchieri d'acqua, 2 per da 1 a 3 bicchieri, 4 per da 3 a 5, ecc. Questi sono i gruppi sui quali può essere misurata la presenza di un andamento. Le variabili di confondimento determinano strati; strati comuni potrebbero essere gruppi per età o per età-sesso. La seguente tabella può essere utilizzata come esempio.

Frequenza di fumatrici di sigarette tra donne con Infarto del Miocardio (MI) e Controlli (C), Stratificati per età

Strato		1		2		3		4		5	
Gruppo		25-29		30-34		35-39		40-44		45-49	
Sig./die	Score	MI	C	MI	C	MI	C	MI	C	MI	C
Nessuno	0	1	131	0	188	3	161	11	169	23	157
1-24	1	1	101	6	152	12	130	21	134	42	97
>24	2	4	51	15	83	22	65	39	68	34	52

Da Schlesselman, JJ. *Case-Control Studies*. Oxford Univ. Press, NY, 1982, p. 205.

In questo esempio si potrebbe inserire lo Strato 1 nella prima schermata, quindi premere <F2> ed inserire lo Strato 2 nella seconda schermata. Dopo aver immesso in tal modo tutti cinque gli strati, è sufficiente premere <F4> per vedere i risultati.

STAMPA O INVIO DEI RISULTATI AD UN FILE

Per inviare i calcoli della schermata corrente alla stampante, premere <F5>. Se si preferisce che ciò che è stato prodotto sia scritto su un file piuttosto che stampato, premere invece <F6>. Appena si preme <F6> viene richiesto un nome per il file da creare. Una volta che il file è aperto, i successivi comandi <F5> aggiungeranno materiale al file anziché inviarlo alla stampante. Premendo ancora <F6> il file viene chiuso. Se si fornisce il nome di un file già esistente con il comando <F6>, il nuovo materiale verrà aggiunto a quello esistente nel file senza danneggiare i suoi contenuti originali.

EPITABLE

EPITABLE, ideato e programmato da Denis Coulombier, è una recente aggiunta a Epi Info. La documentazione è compresa in un file di assistenza fornito con il programma. Per visualizzare il file, premere <F1>. Per conoscere meglio il programma si consiglia di sperimentare direttamente le molteplici funzioni statistiche e i grafici disponibili.

Per ulteriori dettagli si veda il capitolo 15 del Manuale elettronico.

Capitolo 12

VALIDATE: DOPPIA IMMISSIONE DI DATI E VERIFICA

INTRODUZIONE

Con Epi info viene fornito un programma speciale per verificare se due file di dati sono identici e, se non lo sono, mostra tutte le differenze. Questa funzione del programma VALIDATE è utile se si adotta il metodo di immettere *due volte* in ENTER gli *stessi dati*, in modo di essere più sicuri del corretto trasferimento dei questionari cartacei nella base di dati.

COME AVVIARE VALIDATE

Prima di avviare il programma VALIDATE, devono essere disponibili due file .REC contenenti gli stessi dati di Epi Info. Se ogni record contiene un unico numero di identificazione o nome, i record non devono essere necessariamente immessi nei due file nello stesso ordine. Ci dovrà essere uno stesso numero di record nei due file e comunque questi due file devono essere molto simili tra loro. Verranno rilevati solo gli errori di immissione dei dati delle coppie di record che hanno lo stesso numero identificativo (IDNUM).

VALIDATE può essere avviato dal menu PROGRAMMA di EPI6 o dal DOS digitando VALIDATE. Appare una finestra dialogo dove immettere le seguenti informazioni:

- I nomi dei due file Epi Info da mettere a confronto
- L'opzione di indirizzamento dei risultati sullo schermo, a stampa o su file (se i risultati saranno trasmessi a un file sarà necessario immettere il nome di un file di destinazione).
- Se si desidera includere i record cancellati nel confronto (Include deleted records)
- Se i record sono, o potrebbero essere, in un ordine diverso, scegliere "Use Unique IDs" (Usare gli identificativi univoci). Verrà visualizzato un messaggio sullo schermo che vi richiederà di scegliere il campo contenente l'identificatore record univoco. Quando i record contenuti nei due file sono stati immessi nello stesso identico ordine, non è necessario usare il campo identificativo univoco.

Al termine, cliccare su OK o premere <Invio> due volte: il programma esaminerà entrambi i file e produrrà un resoconto delle differenze tra i record in ciascun file e dei campi in cui tali differenze si trovano. A questo punto sarà possibile apportare modifiche tramite ENTER o tramite l'opzione UPDATE di ANALYSIS.

USARE ENTER PER LA CONVALIDA DEI DATI

In alternativa a VALIDATE è possibile usare direttamente ENTER per lo stesso scopo di controllare la corretta immissione dei dati. L'opzione 4 nel menu ENTER, "Reenter and verify records in existing data file" consente all'operatore di reimmettere ogni record e di verificare i dati immessi la prima volta.

Quando "Reenter and verify records in existing data file" viene scelto dal menu ENTER, viene richiesto il nome del campo che identifica in modo univoco ciascun record. Immettere il nome di questo campo identificativo. Quando la verifica è attivata, l'utente immette (o accetta) i dati nel campo identificativo (key) e il programma esegue un'AUTOSEARCH automatica in questo campo. Non sarà possibile immettere i dati per nuovi record.

L'operatore reimmette i dati nello stesso modo con cui li ha immessi la prima volta. La differenza è che ENTER confronta ogni immissione con i record esistenti: quando l'inserimento di un campo corrisponde ai dati del file, il cursore si sposta esattamente come nel caso di nuove immissioni. Quando non corrisponde, appare un messaggio e all'operatore viene data la possibilità di reimmettere il campo o di fare un confronto tra l'inserimento originale e quello nuovo per poi fare una scelta.

I record verificati (immessi due volte) verranno segnalati alla fine con ^. Il numero del record che appare nell'angolo inferiore destro dello schermo verrà segnalato con il simbolo 'check' quando il record è stato verificato.

In ANALYSIS è possibile tenere conto dei record verificati. Una nuova variabile di sistema, RECVERIFIED, viene impostata a Y quando il record è verificato e a N quando non lo è stato. Questo consente di scrivere programmi per gestire, ad esempio, un nuovo file, da trasmettere ad un'agenzia centrale, contenente soltanto i record verificati.

Per ulteriori dettagli si veda il capitolo 16 del Manuale elettronico.

Capitolo 13

EXPORT: PRODURRE FILE PER ALTRI SISTEMI

INTRODUZIONE

I file di dati creati in Epi Info hanno i nomi con estensione .REC. Sono interamente in caratteri "ASCII" (stampabili) e possono essere trasmessi per posta elettronica. Se si preferisce analizzare i dati con altri programmi diversi da Epi Info, il programma EXPORT trasformerà i file .REC in file compatibili con altri software in commercio. Per cambiare il formato di file di dati creati con altri programmi in formato Epi Info, utilizzare il programma IMPORT come descritto al capitolo seguente.

Tipi di file Prodotti da EXPORT

Menu	Estensione	Programma
1 - Systat	SYS	Systat (SYSTAT, Inc., Evanston, IL)
2 - SAS	SAS	SAS (Statistical Analysis System, SAS Institute, Inc., Cary, NC)
3 - Delim	SDF	File delimitato con virgole e virgolette. Tutti i campi sono delimitati da virgolette e separati da virgole. I primi tre record contengono la lunghezza del campo, i dati immessi e i nomi dei campi.
4 - Lotus	WKS	Lotus 1-2-3 (Lotus Development Corp., Cambridge, MA)
5 - SPSS-X	SPS	SPSS (Statistical Package for the Social Sciences, SPSS Inc., Chicago, IL) Questo formato è utilizzabile con la versione per mainframe di SPSS.
6 - EpiStat	EPI	EPISTAT. È un programma di analisi statistica per epidemiologi (Epistat Services, Richardson, TX). Questo formato è utilizzabile con la versione "shareware" originale del programma. Epistat, nella versione commerciale, può essere utilizzato con il formato Fixed-Field descritto di seguito.
7 - dBASE II	DBF	dBASE II (Borland International, Scotts Valley, CA)
8 - dBASE III	DBF	dBASE III e III Plus (Borland International, Scotts Valley, CA)
9 - dBASE IV	DBF	dBASE IV (Borland International, Scotts Valley, CA)
10 - BASIC	DAT	I campi sono delimitati da doppie virgolette e separati da virgole. È identico a DELIM, ma senza informazioni inerenti al nome e alla lunghezza dei campi.
11 - SPSS	SPS	È la versione per microcomputer di SPSS (vedi prima).

12 - Fixed	CAR	Formato a campo fisso (Fixed) o a scheda (card). È un formato diffuso per lo scambio di informazioni, specialmente su computers mainframe.
13 - Statpac	nessuno	StatPac (Walonick Associates, Minneapolis, MN), è un programma statistico per MS-DOS.
14 - MULTLR	DAT	MULTLR, COXPH e KMSURV. Programmi public-domain per regressione logistica, Cox propotional hazard e analisi di sopravvivenza secondo Kaplan-Meier, prodotti da Nelson Capos eda Eduardo Franco. I programmi sono disponibili presso Epidemiology Monitor, Marietta, Georgia, U.S.A.
15 - Egret	BDF	Egret. Un programma commerciale di statistica avanzata prodotto da Statistics and Epidemiology Research Corporation, Seattle, Washington, U.S.A.

UTILIZZO DI EXPORT

EXPORT produce file per ciascuno dei programmi precedentemente elencati. Per avviare EXPORT, scegliere il programma dal menu PROGRAMMI di Epi Info.

Con il mouse si può spostare il cursore e poi selezionare le opzioni premendo il pulsante sinistro. Per lavorare senza mouse, spostare il cursore da un campo all'altro usando il tasto <Tab> (<Maiusc - Tab> per andare in dietro), e selezionare con <Invio>. Tenendo premuto il tasto <Alt> e premendo uno dei caratteri evidenziati si selezionerà direttamente l'opzione.

Digitare il nome del file che si vuole esportare e il formato che desiderate. EXPORT produrrà il nuovo file portandosi poi al menu EPI6 o al DOS. Se si desidera vedere il nuovo file si può utilizzare EPED.

Per ulteriori informazioni sul formato dei tipi di file più diffusi da importare, vedere il Manuale elettronico, capitolo 16, alla voce: Note su particolari tipi di file.

AVVIO DI EXPORT DA DOS

EXPORT può essere avviato direttamente da DOS con i seguenti parametri:

EXPORT <Nome del file di Input>.REC <Nome del file di Output> <Codice numerico di conversione>

I codici numerici utilizzati nella conversione sono i seguenti:

- | | | |
|-----------|--------------|-------------|
| 1. SYSTAT | 6. EPISTAT | 11. SPSS-PC |
| 2. SAS | 7. dBASE II | 12. FIXED |
| 3. DELIM | 8. dBASE III | 13. STATPAC |
| 4. LOTUS | 9. dBASE IV | 14. MULTLR |
| 5. SPPS-X | 10. BASIC | 15. EGRET |

per esempio il seguente comando eseguito direttamente dal DOS:

```
C:\>EXPORT DISEASE DISEASE 11
```

produce un file per SPSS-PC chiamato DISEASE.SPS.

LABORATORIO: ESPORTIAMO UN FILE

Avviamo EXPORT dal menù di EPI6.

Lo schermo che appare visualizza una finestra di dialogo che si può comandare o dal mouse o dalla tastiera. Posizioniamo il cursore nel primo campo e premiamo <Invio> per visualizzare i file .REC disponibili. Premere il tasto <Tab> per posizionare il cursore sull'elenco dei file e poi, usando i tasti freccia, andiamo su OSWEGO.REC e premiamo <Invio> per selezionare questo file.

Con il mouse, possiamo selezionare le opzioni dalla finestra di dialogo o spostare il cursore, con un semplice clic con il pulsante di sinistra, quando il cursore si trova sulla posizione desiderata.

Il cursore si dovrebbe trovare ora nella parte della finestra di dialogo chiamata *Output Format*. Per esportare un file per utilizzarlo in SAS, con il tasto <freccia in basso> o mediante il mouse, spostiamoci su SAS, la seconda voce in elenco. Premiamo ora il pulsante sinistro, oppure <Invio>, per effettuare la selezione.

Nella casella *Output File Name* appare OSWEGO.SAS. Premiamo <Invio> per confermare. Il cursore si trova ora sul pulsante OK. Premiamo <Invio> di nuovo per dare inizio alla conversione.

Per visualizzare il file prodotto possiamo andare in EPED e aprire il file OSWEGO.SAS.

Per ulteriori dettagli, si vedano i capitoli 16 e 33 del Manuale elettronico.

Capitolo 14

IMPORTARE DATI DA ALTRI PROGRAMMI

INTRODUZIONE

IMPORT permette l'utilizzo in Epi Info di file creati con altri sistemi per elaborarli oppure per riconvertirli ancora in altri formati. Il programma accetta file in formato lunghezza fissa (fixed-length) o delimitati da virgole (comma-delimited) con le stringhe o il testo dei campi delimitati da virgolette. Inoltre può importare file Lotus .WKS e .WK1 o i file dBASE II o dBASE III (.DBF) direttamente.

Occorre avere un'esatta descrizione del nome del campo, della lunghezza del campo e del tipo di dati per poter creare un questionario Epi Info e un file .REC che accetti i dati. Le intestazioni e i commenti nel file originale, non immessi nello specifico formato, dovranno essere rimossi modificando il file, prima di importarli in Epi Info.

Prima di importare un file di tipo a lunghezza fissa o delimitato da virgole, bisogna creare un questionario Epi Info nel formato del file da importare. Per esempio, se il primo campo del file è RECORD NUMBER, il primo item del questionario dovrà essere RECORDNO seguito dal numero di caratteri di sottolineatura utilizzati per impostare il campo RECORD NUMBER nel file originale. È più semplice utilizzare caratteri di sottolineatura per tutti i campi a meno che non si sia assolutamente certi che contengano solo numeri. Il questionario non deve essere "elegante". È sufficiente utilizzare frasi semplici per creare nomi appropriati ai campi, seguite da caratteri di sottolineatura sulla destra. L'impaginazione potrà essere più o meno così:

Record No ____ Nome _____ Età ____ Sesso _

oppure:

Record No ____

Nome _____

Età ____

Sesso _

Dopo aver creato il questionario nel programma di scrittura, si deve creare il file .REC corrispondente, con ENTER. Si segue la stessa procedura della creazione di un nuovo file dati a partire da un questionario. Quando il questionario comparirà sullo schermo di ENTER, premere <F10> per uscire dal programma, lasciando vuoto il nuovo file .REC, pronto per ricevere i dati importati. Quando sono pronti sia il file .REC vuoto sia il file di dati da importare, selezionare IMPORT dal menu principale di Epi Info.

L'importazione dei file di dBASE o Lotus 1-2-3 è più semplice, ed è sempre da preferire se il programma da cui si vogliono importare i dati può esportarli in questi formati. I nomi dei campi del file da importare verranno trasferiti automaticamente al file .REC di Epi Info. I fogli elettronici di Lotus da importare devono però portare i nomi dei campi sulla prima riga, e soltanto sulla prima riga che, inoltre, non deve contenere formule.

È possibile eseguire IMPORT con il mouse. Se si adoperava il mouse per spostare il cursore, si selezionano le opzioni premendo il pulsante sinistro. Senza mouse, spostare il cursore da un campo all'altro con il tasto <TAB> (<Maiusc-Tab> per andare in dietro), e

premere <Invio>. Per selezionare l'opzione direttamente, tenete premuto il tasto <Alt> e premere una delle lettere evidenziate.

Inserire i nomi dei file e la scelta del formato dati e poi selezionare il pulsante OK (Alt-K). Quando l'operazione è terminata, IMPORT ritornerà al menu principale. È possibile esaminare i record nel nuovo file attraverso il programma ENTER. Sarà necessario qualche tentativo prima di produrre un questionario che sia perfettamente abbinato alla serie esterna di dati.

AVVIO DI IMPORT DA DOS

IMPORT può essere avviato direttamente da DOS con i seguenti parametri:

IMPORT <Nome del file di input> <Nome del file (.REC) di output> <Codice numerico di conversione>.

I codici numerici utilizzati nella conversione sono i seguenti:

- 1 - Fixed field
- 2 - Comma-delimited
- 3 - Lotus 1-2-3
- 4 - dBase

```
C:\> IMPORT STUDY.DBF STUDY.REC 4
```

produce un file Epi Info .REC da un file dBASE. Se il file importato è del tipo 1 o 2 o se il file Lotus 1-2-3 non riporta i nomi dei campi sulla prima riga, si dovrà seguire la procedura descritta all'inizio del capitolo.

LABORATORIO: IMPORTIAMO UN FILE DI UN ALTRO PROGRAMMA

Creiamo innanzitutto un file da importare, utilizzando il programma EXPORT. Usiamo il file OSWEGO di Epi Info e scegliamo l'opzione per il formato fisso o card-image, nella finestra di EXPORT. Il file risultante sarà chiamato OSWEGO.CAR.

Per dire a IMPORT quali tipi di campi sono presenti nel file OSWEGO.CAR, dobbiamo creare un questionario con la descrizione esatta dei campi e quindi creare un file .REC dal questionario. Fortunatamente esiste già un file OSWEGO.QES sul disco rigido.

1. Avviamo ENTER dal menù principale.
2. Digitiamo IMPORTATO nella casella del nome del nuovo file .REC.
3. Digitiamo "2" come opzione (crea un nuovo file .REC da un questionario).
4. Inseriamo OSWEGO.QES come file questionario da cui creare il file .REC.
5. A questo punto usciamo da ENTER senza inserire alcun dato.

Ora avviamo IMPORT. Impostiamo IMPORTATO come nome per il file Epi Info che riceverà i dati e OSWEGO.CAR per il file da importare. Scegliere l'opzione 1, record di lunghezza fissa, e premere <F4>.

Per provare il nuovo file, avviamo ENTER ed esaminiamo il file di dati IMPORTED. Ora ci dovrebbero essere 75 record nel file. Premere <F7> per curiosare tra i record e confermare che siano corretti.

Quando si importano file dBASE, non è necessario fare prima un questionario Epi Info o un file .REC. IMPORT costruirà il file automaticamente. Ciò è vero anche per Lotus .WKS e per i file .WK1 se la prima linea del foglio elettronico Lotus contiene i nomi dei campi.

Per ulteriori dettagli si vedano i capitoli 17 e 33 del Manuale elettronico.

Capitolo 15

EPIGLUE: UN SISTEMA INFORMATIVO EPIDEMIOLOGICO E DI MENU

INTRODUZIONE

Il menu per Epi Info, Versione 6, è in realtà una copia del programma EPIGLUE.EXE con un nuovo nome. EPIGLUE fornisce gli strumenti per creare menu personalizzati, con menu a tendina e finestre di dialogo creati a scelta tramite l'editing di un file di testo contenente nomi e comandi di menu. EPIGLUE offre inoltre dispositivi di editing e di visualizzazione che forniscono un accesso rapido ai file di testo. Il menu funziona tramite mouse o tastiera, e si può inserire uno sfondo decorativo scelto da un file di caratteri a piacere.

EPIGLUE dà la possibilità di visualizzare report di testo o tabelle, o in formato di testo, o in formato Ipertesto, attraverso il quale si può scorrere cliccando sulle parole chiave evidenziate. È possibile da EPIGLUE chiedere a programmi come Harvard Graphics o Epi Map di visualizzare grafici o mappe. È possibile inoltre aggiungere ai menu funzioni di editing e di gestione dei file, un calendario, una calcolatrice o una tabella ASCII, semplicemente aggiungendo una riga di codice al file .MNU che gestisce il menu.

Tramite EPIGLUE, si può impostare un sistema informativo epidemiologico che visualizza i report più recenti da un sistema di database selezionandoli direttamente da menu. I report possono essere in formato testo, grafico, mappe o altri.

EPIGLUE è la struttura capace di tenere insieme i dati prodotti dai molteplici sistemi in una rete locale LAN (Local Area Network), o avviare programmi da diverse fonti come se fosse un unico sistema.

L'USO DELL' EDITORE TESTI INCORPORATO IN EPI6

Si è già detto che il menu principale di Epi Info 6 non è altro che una copia di EPIGLUE personalizzata. I menu UTILITÀ, FILE e MODIFICA richiamano funzioni incorporate in EPIGLUE e possono essere aggiunte in qualsiasi applicazione creata dall'utente (per esempio, il menu UTILITÀ è stato aggiunto in questa versione italiana di Epi Info 6).

Tramite l'Editore si possono visualizzare e modificare i file tramite finestre. Anche se questo editore testi è comodo perché facilmente accessibile da menu, la dimensione massima dei file che si potranno modificare è di 64 K.

Per questa traduzione italiana di Epi Info, è disponibile una guida in linea contestuale per tutte le opzioni incorporate di EPIGLUE, compreso il menu principale di EPI INFO. Basta quindi premere il tasto <F1> con il cursore posizionato sulla voce di menu della quale si vogliono spiegazioni, per far apparire la finestra del menu con le istruzioni in italiano.

Tale file di guida è disponibile anche per le applicazioni create dall'utente. Basta copiare nella directory della propria applicazione il file EIS_HELP.HLP presente nella directory EPI6.

LA COSTRUZIONE DEL MENU PERSONALIZZATO

Quando si digita EPI6 dal DOS, questo avvia il programma di menu, EPI6.EXE. Il programma cerca istruzioni del menu in un file chiamato EPI6.MNU, le istruzioni sullo sfondo in EPI6.SCR, e le istruzioni sul file guida in EPI6.HLP. Infine legge EPI6.CFG per la configurazione della stampante e le informazioni relative alla porta a cui è collegata. Questi file verranno descritti in dettaglio successivamente.

È possibile avviare EPI6 digitando il nome del file menu di seguito a EPI6:

EPI6 MENU.MNU

In questo modo, EPI6 suppone che il file descrittore del menu si chiami MENU.MNU e che, se esiste un file per lo sfondo, questo si chiami MENU.SCR, il file guida MENU.HLP, e il file della configurazione stampante, MENU.CFG.

Per costruire la propria applicazione è meglio copiare tutti i file EPI6, qualsiasi estensione abbiano, per costruire un nuovo menu, digitando il comando:

COPY EPI6.* MIOMENU.*

dove MIOMENU è il nome della propria applicazione di non più di otto caratteri. Questo comando creerà 5 file nuovi: MIOMENU.EXE, MIOMENU.MNU, MIOMENU.HLP, MIOMENU.SCR, MIOMENU.CFG. È meglio utilizzare COPY, piuttosto che RENAME, per conservare una copia dei file EPI6.* originali.

Ora si possono personalizzare i file MIOMENU.SCR e MIOMENU.MNU per adattarli alle funzioni del nuovo programma.

I FILE DI IMPOSTAZIONE DEI MENU (.MNU)

Un file di impostazione di menu è un file di testo (ASCII) che contiene le informazioni sui menu a tendina che appariranno sullo schermo, sulle voci contenute in ciascuna tendina e sui comandi da associare a ciascuna voce di menu. Contiene inoltre una descrizione associata a ciascuna voce di menu, che apparirà sulla riga di stato del menu ogni volta che il cursore è spostato sulla voce.

Per inserire commenti nel file si utilizza un asterisco (*) come primo carattere della riga: tutto ciò che inizia con un asterisco verrà ignorato dal programma.

Il file è articolato in tre sezioni che si possono presentare in qualsiasi ordine:

- L'impostazione dei menu e delle voci di menu (menu layout)
- I comandi da associare a ciascuna voce di menu
- La descrizione delle voci di menu che appariranno sulla barra di stato

IMPOSTAZIONE DEI MENU E DELLE VOCI DEI MENU

Un menu inizia con il nome del menu, seguito da uno spazio e la parola riservata MENU. La riga dovrebbe essere seguita da una serie di istruzioni che iniziano con BEGIN e terminano con END. In questo blocco, sono consentite quattro istruzioni: MENUITEM, MENUITEM SEPARATOR, MENUEDIT, e POPUP.

MENUIITEM

Sintassi: MENUITEM "&Testo da visualizzare {^ <Simbolo> }", <Blocco comandi>

MENUIITEM rappresenta una voce di un menu a tendina. La stringa di caratteri che si trova tra virgolette e che segue l'istruzione è la voce che apparirà nella tendina. La lettera che segue il simbolo "&" viene evidenziata e diventa la lettera da usare con <ALT> come scorciatoia per eseguire direttamente la voce di menu.

Il ^<Simbolo> (facoltativo) viene usato per aggiungere un simbolo sulla parte destra della voce nel menu a tendina. I caratteri predefiniti in EPI6 sono una freccia (carattere ASCII 16), per indicare che una voce di menu richiama un submenu. Il ^<Simbolo> consente comunque di definire un simbolo personalizzato. Il simbolo potrebbe essere un carattere o, se preceduto dal carattere # e un numero, un codice numerico ASCII. I simboli possono indicare sul menu che il programma o la funzione richiamata della voce accetta parametri. Premendo <Ctrl.Invio> sulla voce è possibile per l'utente inserire i parametri richiesti.

Ad esempio, i simboli di note musicali a fianco delle voci che richiamano i programmi di Epi Info, nel menu di EPI6, stanno ad indicare che premendo <Ctrl-Invio> è possibile inserire il nome del file da aprire con quel programma. Il nome del file è dunque il parametro accettato prima dell'esecuzione del programma, il quale aprirà direttamente il file indicato.

<Blocco comandi>, che segue la virgola, è il nome che verrà usato per un blocco di comandi che verranno eseguiti quando si seleziona la voce (vedi sotto) e anche per visualizzare sulla barra di stato la descrizione della voce quando il cursore si trova sulla voce di menu. Quest'ultima funzione verrà descritta in seguito (vedi il paragrafo Tabella di stringhe).

Ci sono quattro comandi predefiniti che si possono inserire come nomi di blocco di comandi: IDQUIT, IDCALC, IDCLND, IDASCII, e IDSETUP.

Comando	Azione
IDQUIT	Esce dal programma e poi a DOS
IDCLND	Visualizza un calendario che calcola la settimana epidemiologica (come viene usata negli Stati Uniti) per qualsiasi data selezionata.
IDCALC	Visualizza una calcolatrice
IDASCII	Visualizza una tabella di caratteri ASCII e i loro codici
IDSETUP	Visualizza una finestra di dialogo per l'impostazione della directory di default.

Esempi:

MENUIITEM "&Analysis", DoAnalysis

Visualizza "Analysis" come un'opzione di menu, con la lettera "A" evidenziata. Quando si sceglie questa voce, vengono eseguiti i comandi del blocco di comandi DoAnalysis.

MENUIITEM "&Analysis^#14, DoAnalysis

Simile all'esempio precedente, ma visualizza "Analysis", con una nota musicale a destra dell'opzione di menu; questo indica che avvia un programma e che è possibile immettere un parametro (in questo caso un nome di file .REC o .PGM) premendo <Ctrl-Invio> (consultare PARAMETER e FILEPARAMETER sotto).

MENUITEM SEPARATOR

Sintassi: MENUITEM SEPARATOR

MENUITEM SEPARATOR visualizza una riga orizzontale nella tendina del menu. Serve per separare gruppi di funzioni nel menu.

MENUEDIT

Sintassi: MENUEDIT

MENUEDIT aggiunge una serie standard di due submenu, FILE e MODIFICA, che contengono una serie di funzioni per il trattamento e la modifica di testi. Può trovarsi in qualsiasi punto della barra orizzontale del menu.

POPUP

Sintassi: POPUP "&Testo da visualizzare"

```
BEGIN
    MENUITEM          (vedi sopra, una o più voci)
    MENUITEM          (vedi sopra, una o più voci)
END
```

POPUP definisce la barra orizzontale dei menu a tendina. Ogni voce presente sulla prima riga orizzontale dello schermo corrisponde a un menu POPUP. Selezionando una voce, appare la tendina corrispondente al menu, come definita dai MENUITEM contenuti nella definizione del POPUP.

Un menu POPUP è solitamente una voce del menu principale, ma può anche essere inserito fra due MENUITEM, all'interno di un altro menu POPUP. In questo caso si avrà un sottomenu a tendina che si attiverà alla selezione del MENUITEM corrispondente.

Ecco un esempio funzionante di menu e sottomenu:

```
MIOMENU MENU
```

```
BEGIN
```

```
    POPUP "&Utilità"
```

```
        BEGIN
```

```
            MENUITEM "Ca&lendario", IDCLND
```

```
            MENUITEM "&Calcolatrice", IDCALC
```

```
            MENUITEM "Tabella &ASCII", IDASCII
```

```
            MENUITEM "Item &1", AZIONE1
```

```
        MENUEDIT
```

```

MENUITEM "Item &2", AZIONE2
POPUP "&Sottomenu 1"
    BEGIN
        MENUITEM "Item &3", AZIONE3
        MENUITEM "Item &4", AZIONE4
        MENUITEM SEPARATOR
        MENUITEM "Item &5", AZIONE5
    END
MENUITEM "Item &6", AZIONE6
END
END
END

```

In questo esempio verrà creata una barra di menu orizzontale con la voce "Utilità" creata dal menu e le voci "File" e "Modifica", inserite dal comando MENUEDIT. All'interno del menu "Utilità" è presente un sottomenu corrispondente alla voce sottomenu.

Per provare questo menu basta trascriverlo in un file di testo, assegnargli il nome Miomenu.mnu e salvarlo in c:\epi6. Per avviarlo digitare il seguente comando:

EPIGLUE MIOMENU.MNU

Il programma EPIGLUE caricherà il file appena creato per impostare i propri menu invece del file EPIGLUE.MNU predefinito.

LA TABELLA DI STRINGHE

La Tabella di Stringhe fornisce una guida di frasi brevi che appaiono in fondo allo schermo per spiegare all'utente la funzione eseguita selezionando la voce di menu. Appaiono automaticamente quando il cursore a barra si trova sulla voce di menu corrispondente.

L'impostazione delle stringhe da visualizzare viene eseguita tramite il comando:

STRINGTABLE

Sintassi:

```

STRINGTABLE
BEGIN
    <Nome del blocco di comando> "<frase di guida>"
END

```

STRINGTABLE è un elenco di frasi o "stringhe" che corrispondono alle varie opzioni di menu comprese nel blocco BEGIN - END, e immediatamente successive all'istruzione STRINGTABLE. Quando una stringa viene definita per un dato blocco di comando, la stringa appare in fondo alla schermata del menu quando il cursore è posizionato su quella voce. Serve a fornire brevi frasi di guida che danno informazioni sulla natura della scelta di menu. Se una determinata voce non ha una frase di descrizione, la riga di stato in fondo allo schermo rimane vuota.

STRINGTABLE è facoltativo.

Esempio:

```
STRINGTABLE
  BEGIN
    AZIONE 1  "Esegue azione 1"
    AZIONE 2  "Esegue azione 2"
    AZIONE 3  "Esegue azione 3"
    AZIONE 4  "Esegue azione 4"
    AZIONE 5  "Esegue azione 5"
    AZIONE 6  "Esegue azione 6"
  END
```

I COMANDI DI BLOCCO

La parte rimanente del file .MNU serve a definire i comandi da eseguire per ciascuna opzione di menu. Il blocco di istruzioni tra BEGIN e END è simile a un file batch del DOS. Qualsiasi istruzione che appare verrà eseguita esattamente come se si digitassero le istruzioni dal DOS. Si può chiamare un programma esterno come ANALYSIS, e usare quasi tutti gli altri comandi dei file batch del DOS.

FILEDIALOG, DIALOG, VIEW, e HELP sono comandi aggiuntivi che vengono usati in blocchi di comando per creare finestre di dialogo interattive per ottenere dati dall'utente, visualizzare file in una finestra editor/viewer e fornire le funzioni hypertext. Per ulteriori istruzioni su questi comandi consultare il manuale elettronico in linea.

FILE GUIDA IPERTESTUALI

Gli ipertesti non sono altro che testi con alcune parole evidenziate. Cliccando su quelle parole appare un'altra finestra contenente ulteriori informazioni su quell'argomento e altre parole evidenziate, rimandanti a loro volta ad altri argomenti. Questo intreccio di informazioni a portata di mouse è un ipertesto, e EPIGLUE permette di costruire tali ipertesti per essere utilizzati come file di guida. Il manuale elettronico in linea non è altro che una guida ipertestuale di EPIGLUE.

La creazione di guide ipertestuali da agganciare alla propria applicazione richiede un certo lavoro e una certa esperienza. Si rimanda pertanto al manuale elettronico in linea per le informazioni sulla sintassi degli ipertesti e sulle modalità di aggancio alle applicazioni create dall'utente.

LABORATORIO: CURIOSIAMO IN EPIGLUE

Avviamo innanzitutto EPIGLUE dal menu ESEMPI di Epi Info 6. La prima cosa da fare è esaminare le opzioni di tutti i menu. Si noterà che nei menu non ci sono funzioni particolari ma soltanto esempi su come sia possibile personalizzare una maschera di menu per adattarla alla propria applicazione.

Premendo <F1> in qualsiasi momento, appare la spiegazione della funzione selezionata in italiano. Il lavoro da compiere è allora quella di selezionare ogni opzione e vedere cosa succede, aiutandosi con il file della guida.

Una volta curiosato fra le varie funzioni fra le funzioni si può, passare alla creazione della propria applicazione. La logica di EPIGLUE è quella di offrire tutta la gamma di possibilità di costruzione di menu, e le sue opzioni non sono che esempi di menu. Quando si trova una funzione interessante, basta aprire il file EPIGLUE.MNU e copiare in MIOMENU.MNU (la propria applicazione) il blocco di comandi che interessa.

Nel menu UTILITÀ ci sono due funzioni per la creazione automatica di menu a tendina e dello sfondo dell'applicazione. Avviare entrambi i programmi e seguire le istruzioni a video e ipertestuali.

Per ulteriori dettagli, si vedano i capitoli 20 e 37 del Manuale elettronico.

Capitolo 16

SINTESI DEI COMANDI DI EPI INFO

COMANDI DI EPED

Tasti funzionali <ctrl>

Diverse funzioni di uso frequente sono accessibili tramite i 10 tasti funzionali oppure i 10 <Ctrl> + tasto funzionale. Le funzioni sono:

<Ctrl F1 >	Inizio blocco
<Ctrl F2 >	Fine blocco
<Ctrl F3 >	Copia blocco
<Ctrl F4 >	Muovi blocco
<Ctrl F5 >	Cancella blocco
<Ctrl F6 >	Trova testo
<Ctrl F7 >	Sostituisci testo
<Ctrl F8 >	Giustifica paragrafo
<Ctrl F9 >	Modo - cambia il modo da Word rap a Text (TXT) e viceversa

Tasto <scrl lock>

Premendo il taso Scroll Lock (Scrl Lock) come un interruttore, si accenderà una spia luminosa sulla tastiera. Durante l'attività del tasto Scrl Lock alcuni tasti assumono una funzione particolare:

<Tasti direzionali (frecche)>	Tracciano sullo schermo un riquadro (box)
<Tasto Home o \curvearrowright >	Alterna il box a linea singola o doppia
<Tasto Fine>	Cancella o traccia vari caratteri ombreggiati
<Tasti lettere tastiera>	Simboli alfabeto greco

<F1> il comando help (aiuto)

<F2> i comandi file

OPEN FILE THIS WINDOW <Ctrl K O>C (Control K O, non "zero") "Apri un file in questa finestra"
Apri un file nuovo o già esistente. Se non conoscete il nome del file, premete il tasto <INVIO> per ottenere una lista di file. Il segno "*" elencherà tutti i file nella directory.

ADD FILE OTHER WINDOW <Ctrl O A> "Aggiungi un file in un'altra finestra"

Apri un altro file in una "finestra" diversa. Il numero di finestra è indicato nell'angolo inferiore sinistro dello schermo. EPED può gestire quattro finestre contemporaneamente.

CHANGE WINDOW <Ctrl J W> "Cambia finestra"

Se è aperta più di una finestra, mostra un elenco delle finestre aperte.

CLOSE FILE THIS WINDOW <Ctrl O Q> "Chiudi questa finestra"

SAVE FILE TO.. <Ctrl K N> "Salva il file con il nome..."

FILE DIRECTORY "Elenco dei file"

Visualizza un elenco di tutti i file disponibili nella directory corrente.

USE DOS <Ctrl J O> "Usa il DOS"

Questo comando permette di richiamare i comandi del DOS o di accedere ad altri programmi dall'interno di EPED. Tornate ad EPED digitando EXIT.

<F3> il comando EPIAID

Questo comando avvia il menù EPIAID. L'uso di EPIAID è descritto nei Capitoli 21, 28 e 34 nel manuale elettronico dal "menu Manuale".

<F4> i comandi di testo

QUESTIONS <Ctrl Q Q> "Domande"

Questo comando visualizza un menù sovrapposto contenenti i tipi di campo di dati necessari , per creare un questionario Epi Info. I campi possono essere immessi nel questionario posizionando il cursore nel punto di immissione dati del questionario e quindi selezionando il tipo di campo di dati dal menù utilizzando la barra selezionatrice.

CENTER CURRENT LINE <Ctrl O C> "Centra nella linea corrente"

<F7> i comandi di ricerca

FIND <Ctrl-F6> o <Ctrl Q F> "Cerca"

Ricerca stringhe fino a 67 caratteri. Una finestra di messaggi mostra l'ultima stringa di ricerca inserita.

C= Corrispondenza esatta rispetto alle lettere maiuscole e minuscole. U = Ricerca verso l'inizio del documento del file partendo dalla posizione corrente del cursore. W=Word (parola) Trova la stringa soltanto quando circondata da spazi. B= Trova soltanto all'interno di un blocco di testo che è stato selezionato con i comandi di blocco.

FIND AND REPLACE <Ctrl-F7> o <Ctrl Q A> "Cerca e sostituisci"

Cerca una stringa e la sostituisce con un'altra. Una finestra di messaggi mostra l'ultima ricerca e rimpiazza le stringhe inserite. G=Cerca in tutto il file (globalmente). N=Sostituisce senza conferma (no ask - non chiede conferma). 1...n = Effettua il numero di cambiamenti specificato.

<F8> il BLK o comandi di blocco

BEGIN BLOCK <Ctrl F1> o <Ctrl K B>	"Inizio blocco"
END BLOCK <Ctrl F2> o <Ctrl K K>	"Fine blocco"
HIDE BLOCK DISPLAY <Ctrl K H>	"Evidenziazione del blocco"
COPY BLOCK TO HERE <Ctrl F3> o <Ctrl K C>	"Copia blocco qui"
MOVE BLOCK TO HERE <Ctrl F4> o <Ctrl K V>	"Muove blocco qui"
DELETE BLOCK <Ctrl F5> o <Ctrl K Y>	"Cancella blocco"
WRITE BLOCK TO FILE <Ctrl K W>	"Scrivi blocco nel file"
READ FILE IN AS BLOCK <Ctrl K R>	"Leggi file come blocco"

<F9> il comando save

Tale comando salva il file. Se non è stato designato nessun file, prima invita a fornire un nome di file.

<F10> il comando done

Questo comando lascia EPED e ritorna al menù di EPI o al DOS.

Altri comandi a tastiera

Inverte minuscolo-maiuscolo	Ctrl-O O	Linea superiore:	Ctrl-E
Minuscolo	Ctrl-O V		o Freccia in su
Maiuscolo	Ctrl-O U	Linea inferiore:	Ctrl-X
Evidenzia una singola parola	Ctrl-K T		o Freccia in giù
Precedente posizione cursore	Ctrl-Q P	Scorri su	Ctrl-W
A capo come la riga sopra	Ctrl-J B	Scorri giù	Ctrl-Z
A capo come la riga sotto	Ctrl-J E	Pagina superiore	PgUp o Ctrl-R
Prossimo paragrafo	Ctrl-Q Z	Pagina inferiore	PgDn o Ctrl-C
Precedente paragrafo	Ctrl-Q W	Inizio linea	Home o Ctrl-Q S
Carattere sinistro:	Ctrl-S	Fine linea	End o Ctrl-Q D
	o Freccia a sinistra	Tabulatore	Tab o Ctrl-I
Carattere destro:	Ctrl-D	Tabulazione inversa	Maiusc-Tab
	o Freccia a destra	Sommità schermo	Ctrl-Home Ctrl-Q E
Parola sinistra:	Ctrl-A	Base schermo	Ctrl-End o Ctrl-Q X
	o Ctrl-freccia a sinistra	Nuova linea	Enter o Ctrl-M
Parola destra:	Ctrl-F	Inserisci riga	Ctrl-N
	o Ctrl-freccia a destra	Cancella carattere corrente	Del o Ctrl-G

Cancella carattere a sinistra	Backspace o Ctrl-H	Cancella fino a fine linea	Ctrl-Q Y
Cancella parola	Ctrl-T	Cancella linea	Ctrl-Y
		Disattiva menù	Esc

COMANDI DI ANALYSIS

BAR

Sintassi: BAR <Nome della variabile> {/Y=<min>-<max>} {/SUM = <Nome del campo>}
{/LABEL=<Testo>}

Il comando BAR costituirà un grafico a barre della variabile specificata. È utile per visualizzare i dati relativi ai valori di variabili sia qualitative che quantitative. Normalmente, le barre verticali rappresentano conteggi dei record con i singoli valori della variabile. Se viene specificato SUM, le barre rappresentano le somme della variabile immessa nel parametro SUM (variabili quantitative).

I grafici a barre differiscono dagli istogrammi in quanto le barre non sono separate da spazi e vengono omessi i conteggi relativi a valori nulli.

Normalmente l'etichetta sull'asse verticale è "Count". Se il parametro /LABEL=<Testo> è fornito, l'etichetta usata è quella data come Testo.

BREAK

Sintassi: BREAK <nome della variabile> {<nome della variabile> ...ecc.}

BREAK realizza il comando contenuto in BREAKCMD ogni volta che il valore del <nome della variabile> cambia durante un comando LIST. I record devono essere per prima cosa ordinati sulla variabile specificata. BREAKCMD è normalmente impostato per saltare ad una nuova pagina se il risultato viene inviato alla stampante o al file. BREAKCMD può essere mutato in un appropriato comando di Epi Info utilizzando SET BREAKCMD = (Comando). Consultate SET per ulteriori dettagli.

BROWSE (<F4>)

Sintassi: BROWSE {<lista dei nomi dei campi>}

Il comando BROWSE mostra il file di dati corrente ed ogni file correlato (RELATE), in un formato simile ad un "foglio elettronico". Potete usare i tasti freccia per muovervi e vedere l'intero insieme di dati. In BROWSE, sono disponibili le seguenti funzioni:

<Home><End> Questi tasti portano al primo o all'ultimo record.

<F4> mostra alternativamente i record dell'insieme di dati nel formato questionario, come appaiono nel programma ENTER, o nel formato "foglio elettronico" come quando viene attivato BROWSE.

CLOSE

Sintassi: CLOSE {<nome del file>}

CLOSE chiude il file specificato. Se il file è stato correlato, è equivalente all'estinzione della correlazione.

CLS

Sintassi: CLS

CLS pulisce lo schermo.

CMD

Si potrà utilizzare il comando CMD per definire un blocco di comandi che si comporteranno come un'unica entità, simile ad una subroutine o procedura in altri linguaggi.

Sintassi: CMD <Comando>

<Blocchi di Comandi>

END

Il comando CMD serve anche a avviare un programma esterno da ANALYSIS e a rinviare i risultati ad ANALYSIS. Sintassi: <CMD <nome del comando> {FILE = nome del file .exe} {TABLE} {NOPARAM} {MAXFIELDS=[numero massimo di campi]} {NUMERIC=[numeric field set]}

Questo uso di CMD è descritto nel capitolo 39 del manuale elettronico.

COMBINE

Sintassi:

COMBINE{NAME=<Nome della Variabile>} VALUE=<Valore della Variabile>
<Variabile> {=code} <Variabile> <Variabile>

COMBINE{NAME=<Nome della Variabile>} VALUE=<Valore della Variabile>
<Variabile> TO <Variabile>

COMBINE fa sì che i contenuti di più campi vengano trattati come se fossero nello stesso campo. Vedere a pag. 79 o il manuale elettronico in linea per maggiori dettagli (capitolo 36).

COMMANDS (<F2>)

Premendo il tasto <F2> vi visualizzerà un menu con la lista dei comandi. Si può scegliere il comando desiderato spostando il cursore a barra e poi premendo <Enter>. Il comando verrà automaticamente inserito prima dei nomi delle variabili anche se scelte in precedenza.

COPY

Sintassi: COPY <file d'origine> <file di destinazione>

Copia un file come il comando COPY del DOS, utilizzando la stessa sintassi.

Esempi: COPY A:ALPHA.BAT B:ALPHA.BAT
COPY \EPI6\GAMMA.BAT \DOS\GAMMA.BAT

DEFINE

Sintassi: DEFINE <nome della variabile> _____ {GLOBAL} {CUMULATIVE} <Valore>

```
###  
####.##  
<mm/dd/yy>  
<dd/mm/yy>
```

DEFINE permette la creazione di nuove variabili da utilizzare in ANALYSIS. Vengono forniti il formato e la lunghezza della variabile, come appaiono nei questionari, caratterizzati da caratteri di sottolineatura, simboli di numero (#) o da formati di date americane od europee. La lunghezza dei campi di sottolineatura o numerici può variare come avviene per i questionari.

Vi sono tre tipi di variabili definite -- standard, CUMULATIVE e GLOBAL. La differenza sta nel momento in cui la variabile viene reimpostata al valore mancante o "iniziale".

Le variabili standard vengono reimpostate dopo che ogni record viene elaborato durante una procedura di *FREQ*, *LIST*, o *TABLES*. Esse sono utili per calcoli che devono essere realizzati all'interno di ogni record.

Le variabili CUMULATIVE vengono reimpostate in base al valore mancante o "iniziale" ogni qual volta un nuovo comando che determini la lettura (*READ*) del file viene proposto. Sia le variabili standard che quelle CUMULATIVE vengono reimpostate dal comando *READ* e sono quindi inutili per il mantenimento dei valori che verranno usati durante una serie di programmi con differenti file di dati. Le variabili GLOBAL sono fornite proprio a questo scopo. Esse non sono reimpostate dal comando di lettura (*READ*), e solamente un comando *QUIT* od una esplicita istruzione reimposteranno il loro valore. Se l'utente inserisce la settimana relativa all'elaborazione desiderata, ciò può essere conservato in una variabile GLOBAL in modo che l'utente non sia sollecitato nuovamente ad inserire il numero della settimana.

DOS (<F9>)

Sintassi: DOS {<Comando DOS o programma da avviare>}

Il comando DOS o il tasto <F9> fornisce accesso al DOS da ANALYSIS. Premendo il tasto <F9> o digitando DOS verrà visualizzato il messaggio di richiesta del DOS esattamente come se foste usciti da ANALYSIS. Esce al messaggio di richiesta del DOS e vi consente di dare i comandi del DOS. Per ritornare ad ANALYSIS, digitate "EXIT".

ECHO

Sintassi: ECHO <testo>

ECHO scrive <testo> sullo schermo, ma non alla stampante o ad un file. Questo fornisce un mezzo per comunicare all'utente cosa sta facendo un file attivato automaticamente.

ERASE

Sintassi: ERASE <nome del file>

ERASE cancellerà il file <Nome del file> dalla directory corrente, o, se una directory fa parte del nome del file, da qualsiasi directory.

FOOTER

Sintassi: FOOTER <Numero a piè pagina 1-3> {"<testo a piè pagina>"}

Quelle denominate footer (piè pagina) sono linee di testo che appariranno alla base di ciascuna pagina dei risultati di ANALYSIS. Possono essere specificate fino a tre righe di piè di pagina.

FORM

Sintassi: FORM {/C}

Il comando FORM fornisce un modo per stampare moduli contenenti i dati da ANALYSIS. Il comando stesso fa stampare un questionario che contiene i dati del record corrente. Si potrà unire ai comandi IF, SELECT e PROCESS per stampare uno o tutti i record di un file in formato di questionario.

Normalmente ciascun modulo inizierà in una pagina nuova. Se non desiderate far uscire l'equivalente di NEWPAGE alla fine di ciascun questionario, il parametro /C (per Continua) farà sì che venga saltata una riga prima della stampa del questionario successivo.

FREQ

Sintassi: FREQ <nome di variabile o "*" > {/SUM=<nome di variabile>} {/C}

FREQ produce una tabella dal file specificato nel comando READ, contando i record di ciascun valore della variabile. I record possono essere inclusi o esclusi dal conteggio con il comando SELECT. Quelli indicati come cancellati in ENTER verranno trattati in accordo con l'attuale configurazione del comando SET PROCESS. Se viene fornito più di un nome di variabile, FREQ creerà una tabella separata per ciascuna variabile.

Se viene specificato "/SUM", i risultati rappresentano somme dei valori della variabile fornita nel parametro "/SUM". Se "/C" viene specificata e la variabile è numerica, verranno generati i limiti di confidenza delle proporzioni binomiali. Per ulteriori informazioni relative ai limiti di confidenza consultate il capitolo 32.

<F10>

Consultate QUIT. <F10> e QUIT sono sinonimi per "Abbandonare ANALYSIS e ritornare al menù o al DOS."

GENERATE

Sintassi: GENERATE <Numero di record>

Il comando GENERATE combinato con le funzioni RND, RAN e RECODE produrrà una serie di dati, fino ad un massimo di 32.000 record, in base alle specifiche fornite, ma derivati da numeri casuali. File di questo tipo sono utili per proiezioni, esperimenti, dimostrazioni od esercizi. Diamo per scontato che gli utenti di Epi Info non rischieranno la loro reputazione scientifica o la salute di altre persone, presentando tali dati come risultati reali.

GOTO

Sintassi: GOTO <Nome di etichetta>

GOTO fa sì che il programma riprenda l'esecuzione in corrispondenza dell'etichetta specificata. Le etichette sono indicate dal segno due punti nella prima colonna, come in:

:DOQUESTION

L'etichetta può essere sia prima che dopo l'iterazione GOTO, consentendo così salti indietro o in avanti o movimenti circolari usando le iterazioni IF per controllare il numero delle volte in cui il circolo viene percorso. Sottopercorsi simili a questi, nel BASIC possono essere chiamati con GOTO in combinazione con RETURN alla fine del sottopercorso.

HEADER

Sintassi: HEADER <numero d'intestazione 1-3> {"<testo dell'intestazione>"}

Le linee di intestazione (header) sono quelle che appaiono alla sommità di ciascuna pagina dei risultati di ANALYSIS. Possono essere specificate fino a tre righe di intestazione.

HISTOGRAM

Sintassi: HISTOGRAM <Nome variabile> {/Y=<min>-<max>} {/SUM = <Nome di campo>}

{/LABEL=<Testo>} (Deve essere scritto tutto su una riga oppure su più righe utilizzando il simbolo “\”)

Il comando HISTOGRAM realizzerà un istogramma della variabile specificata. Normalmente, le barre verticali rappresentano conteggi di record con valori differenti della variabile. Se è specificata l'operazione di somma (SUM), le barre rappresentano le somme della variabile specificata col parametro SUM.

Gli istogrammi sono utili per visualizzare i numeri relativi ai casi che si manifestano durante un certo periodo di tempo e questo tipo di grafico viene spesso denominato curva epidemica.

Gli istogrammi differiscono dai grafici a barra in quanto la loro scala ha valori continui, non separati da spazi, ed in quanto visualizzano conteggi anche per valori per i quali la somma è zero (ad esempio nelle curve epidemiche in cui notoriamente l'ascissa è una scala temporale, HISTOGRAM manterrà la sequenza dei giorni anche per quei giorni in cui non ci sono casi).

IF

Sintassi: IF <Condizione(i)> THEN <Conseguenza1> {ELSE <Conseguenza2>}

IF definisce le condizioni ed una o più conseguenze che risultano per ciascun record in cui si verificano le condizioni. Una conseguenza alternativa può essere data dopo un'iterazione ELSE, da realizzarsi se la prima serie di condizioni non è vera. La conseguenza può essere un'assegnazione di valori, simile ad un'iterazione LET (ma senza LET). Non ci può essere un'altra iterazione IF.

Esempi: IF AGE > 15 THEN GROUP = “ADULT” ELSE GROUP =\
“CHILD”

(Notate che “LET” non è incluso nell'iterazione e che “\” viene usato per indicare la continuazione di una linea lunga in un file di programma, ma non sullo schermo).

IMMEDIATE

Sintassi: IMMEDIATE <Comando>

dove <Comando> può essere un comando LET (assegnazione), IF o WRITE.

Il comando IMMEDIATE viene utilizzato con comandi che sono normalmente attivati per ciascun record nel file. Ciò determina l'esecuzione del comando solo una volta nel momento in cui appare in un file di programma o viene digitato sullo schermo, anziché alla fine.

LET (ASSEGNAZIONE DI VALORI ALLA VARIABILE)

Sintassi: {LET} <variabile> = <espressione>

La parola LET è facoltativa.

LET assegna il valore di un'espressione ad una variabile. Se la variabile non è stata definita, ANALYSIS definirà automaticamente una variabile, ma creerà una variabile numerica. È preferibile definire esplicitamente le vostre variabili usando il comando DEFINE.

Il valore verrà calcolato ogni volta in cui un record che soddisfa la condizione di SELECT verrà letto dal file di dati.

LINE

Sintassi: LINE <nome di variabile> {<Nome di variabile>}/SUM = <Nome di campo>}

{/LABEL=<Testo>}

Il comando LINE creerà per prima cosa un grafico a linee della variabile specificata. Normalmente, i punti sulla linea rappresentano i conteggi relativi a record con differenti valori della variabile. Se è specificata "/SUM", i punti rappresentano somme della variabile specificata nel parametro /SUM.

LIST

Sintassi: LIST {<Variabile 1> <Variabile 2> ... <Variabile n>}, oppure LIST {NOT <Variabile 1> <Variabile 2>...ecc.}, oppure LIST * {NOT <Variabile 1> <Variabile 2>...ecc.}

LIST produce una lista elencando l'attuale serie di dati. Se non è specificata alcuna variabile, elenca quante più variabili possibili in un'unica pagina. I numeri dei record vengono elencati automaticamente alla prima voce di ogni riga, ma questa caratteristica può essere eliminata dando il comando SET LISTREC = OFF.

LOAD

Sintassi: LOAD <Nome di file>

Carica un file di programma nella memoria, pronto per l'uso, ma non lo avvia (RUN).

MATCH

Sintassi: MATCH <Variabile ID> <Variabile1> <Variabile2> ecc.

MATCH realizza un'analisi appaiata per produrre odds ratio grezzi, odds ratio con dati appaiati, i limiti di confidenza al 95% secondo di Robins, Greenland e Breslow per odds ratio appaiati e il chi quadrato di Mantel-Haenszel ed il valore p. Il chi quadrato è equivalente al chi quadrato di McNemar se i dati sono trattati come appaiati.

MATCHED

Sintassi: MATCHED

È un miglioramento del comando MATCH. Vi verrà richiesto il nome del file e dopo verrà visualizzato come eseguire il comando. I risultati comprendono quelli per la corrispondenza (match), ma anche una tabella che riporta il numero di serie concordanti e discordanti.

MEANS

Sintassi: MEANS <Variabile numerica da analizzare> <Variabile per raggruppamento> {/N}

Il comando MEANS richiede 2 variabili: il nome di una variabile numerica contenente i dati da analizzare ed il nome di una variabile qualitativa o per categorie da raggruppare. Normalmente MEANS produce prima una tabella che mostra i dati e quindi calcola le statistiche appropriate. Se il parametro /N è incluso nel comando, non viene prodotta nessuna tabella ma vengono mostrate solo le statistiche.

Nella versione Epi Info 6 è possibile indicare solo una variabile quantitativa per ottenerne le statistiche di distribuzione.

MEANS fornisce per dati *distribuiti normalmente*:

ANOVA (per due o più campioni)

Test t di Student (per due campioni)

per dati *distribuiti non normalmente*

Analisi della varianza ad una via di Kruskal-Wallis (per due o più campioni)

Test U di Mann-Whitney = Test di somma dei ranghi di Wilcoxon (per due campioni)

Ulteriori dettagli sono forniti nel capitolo 9 e nel capitolo 32 del manuale elettronico.

NEWPAGE

Sintassi: NEWPAGE

Inserisce, nell'output della stampante o in un file, un numero sufficiente di righe vuote per saltare all'inizio della successiva pagina intera. NEWPAGE è il settaggio di default per break, come se fosse stato dato il comando SET BREAKCMD = NEWPAGE.

OPERATORI

ANALYSIS fornisce i seguenti operatori da utilizzare nelle espressioni aritmetiche e logiche.

Operatore	Funzione
+	Addizione
-	Sottrazione
*	Moltiplicazione
/	Divisione
DIV	Divisione, con risultati interi. Viene scartato ogni resto minore di 1,0.
^	Elevazione a potenza. Eleva un numero alla potenza indicata. Le radici possono essere estratte da valori frazionali come 0,5 per la radice quadrata.
(Insieme di espressioni
)	Insieme di espressioni
>	Maggiore di
<	Minore di
>=	Maggiore o uguale a
<=	Minore o uguale a
=	Uguale a
<>	Diverso da
[Inizio dell'informazione di substringa. Consultate il comando LET.
]	Termine dell'informazione di substringa.
RND (x)	Produce un numero reale casuale compreso tra 0 e il valore di x.
RAN (x)	Produce un numero intero casuale compreso tra 0 e il valore di x.
LOG (x)	Produce il logaritmo in base 10 di x.
LN (x)	Produce il logaritmo naturale (base 2.7182818 = e) di x.
EXP (x)	Innalza la base dei logaritmi naturali, e, alla potenza x.
TRIM (a)	Elimina gli spazi vuoti non necessari da un campo di testo.
TRUNC (x)	Elimina i decimali, generando un numero intero, (arrotonda per difetto)
UND (x)	Arrotonda un numero al valore intero più prossimo, arrotondando per eccesso le frazioni uguali o superiori a 0,5
AND	E logico. Se tutte le condizioni che stanno per essere aggiunte (AND) sono vere, allora l'espressione è vera.
OR	O logico. Se una o più tra le condizioni che stanno per essere poste in alternativa sono vere, allora l'espressione è vera.
NOT	No logico. Se la condizione è falsa, allora l'espressione è vera. Se la condizione è vera allora l'espressione è falsa.

OUTPUT

Sintassi: OUTPUT TABLES | FREQ | SUMTABLES | SUMFREQ <Uno o più nomi di variabili>

Il comando OUTPUT, quando usato con i comandi FREQ, TABLES, SUMFREQ, o SUMTABLES, invia la tabella o la distribuzione di frequenza alla destinazione specificata in un precedente comando ROUTE. Inoltre i comandi:

ROUTE SUMMARY.REC

OUTPUT TABLES DISEASE COUNTY

creeranno un file chiamato SUMMARY.REC, che conterrà dei record che riassumeranno il numero di casi di DISEASE (malattie) per COUNTY. Il file sarà un file .REC standard di Epi Info e conterrà le seguenti variabili:

DISEASE COUNTY COUNT

La variabile COUNT, aggiunta dal programma, conterrà i valori della cella nella tabella che si sarebbe dovuta produrre con

TABLES DISEASE COUNTY

Il comando OUTPUT può essere usato per immagazzinare risultati periodici come i record riassuntivi in un file. Il file riassuntivo può essere elaborato successivamente con ANALYSIS per produrre somme definitive senza rielaborare l'intero file di dati originali. I comandi SUMFREQ e SUMTABLES sono incaricati di elaborare file di dati creati con il comando OUTPUT.

PIE

Sintassi: PIE <Nome della variabile> {SUM = <Nome del campo>}

Il comando PIE realizzerà un grafico a torta (PIE) della variabile specificata. Normalmente, i segmenti della torta rappresenteranno conteggi di record con differenti valori della variabile.

Se è specificata la somma (SUM), i segmenti rappresenteranno somme delle variabili fornite nel parametro SUM.

PROCESS

Sintassi: PROCESS

Questo comando fa sì che ANALYSIS legga il file corrente ed esegua tutti i calcoli senza la produzione di una tabella. Se sarà necessario, l'emissione di dati (l'output) dovrà essere programmata tramite TYPE o comandi simili. È utile per impostare il valore delle variabili cumulative o globali che richiedono la lettura del file per impostare il loro valore.

QUIT (<F10>)

Sintassi: QUIT

Digitando QUIT si esce dal programma ANALYSIS e si ritorna al menù di Epi o al DOS se ANALYSIS era stato avviato direttamente dal DOS. La pressione del tasto di funzione <F10> ha lo stesso effetto.

READ

Sintassi: READ <nome del file di dati> {<nome del file di dati>...fino a 10 nomi}

READ legge un file .REC rendendolo il file di dati attivo. Se non è presente alcun suffisso in un nome di file, aggiunge ".REC" al nome di file prima di tentare di aprire il file. Se non è fornito alcun nome di file, appare una lista di file .REC disponibili. READ *.DBF cercherà i file con estensione .DBF per i file di dBASE. Se viene specificato più di un file, i file devono essere di uno stesso formato, in tal modo file simili con dati differenti vengono letti contemporaneamente (es. archivi di anni diversi).

RECODE

Sintassi: RECODE <Var1> {TO <Var2> {BY <Num>}} Codici

I codici sono rappresentati nella forma:

<Vecchio valore>= <Nuovo valore> <Vecchio valore>=<Nuovo valore> ecc.

ELSE=<Nuovo valore>

Il comando RECODE può essere usato per raggruppare variabili, tradurre codici da un sistema ad un altro o assegnare etichette a codici numerici. La prima variabile elencata è la fonte dei dati da ricodificare. La seconda variabile è il contenitore della nuova informazione ricodificata. Le direttive per la ricodifica sono fornite dopo le due variabili.

REGRESS

Sintassi: REGRESS <variabile dipendente> = <variabile indipendente1> <variabile indipendente2> ecc....

Questo comando realizza regressioni lineari multiple con più di 15 variabili.

RELATE

Sintassi: RELATE <Comune identificatore del nome di variabile> <nome del file>

Il comando RELATE lega uno o più file a quello principale durante l'analisi, utilizzando un identificatore comune per cercare record accoppiati. In un certo senso, realizza le

funzioni del programma MERGE in modo interattivo all'interno di ANALYSIS. (Vedi manuale elettronico capitolo 25 e capitolo 5 di questo manuale.)

RENAME

Sintassi: RENAME <vecchio nome del file> <nuovo nome del file>

RENAME vi consente di rinominare un file sul disco. Se non esiste nessun file con un <vecchio nome di file>, o se già è presente un file con il nome <nuovo nome di file>, il comando fornisce un segnale di attenzione e non agisce.

REPORT

Sintassi: REPORT <nome di file di report>

Il comando REPORT ricava istruzioni da un file di testo creato con EPED e costruisce un report secondo il formato specificato.

I comandi sono preceduti dal simbolo #. All'interno di un report file, sono consentiti i seguenti comandi:

#USES <nome di variabile1> {<nome di variabile2> <nome di variabile3>, ecc.}

Il comando #USES deve trovarsi sulla prima linea di ogni file .RPT.

Ciò informa ANALYSIS della necessità di costruire in memoria una tabella utilizzando le variabili specificate. Tutte le celle ed i totali marginali nella tabella diventano allora disponibili per la stampa come specificato nelle iterazioni fra le parentesi quadre. Le istruzioni che darete tra parentesi quadre sono dei mezzi per collocare qualunque cella o totale marginale della tabella in qualunque punto della pagina, o per scegliere di ometterli dall'output.

#USESSUM <nome di variabile1> {<nome di variabile2> <nome di variabile3>, ecc.}

Questo comando è simile al comando #USES eccetto che la somma dei valori della variabile rimpiazza il calcolo. I due comandi sono analoghi ai comandi TABLES e SUMTABLES in ANALYSIS.

Il testo viene semplicemente stampato come è scritto nella posizione della pagina dove appare. Le espressioni \c e @<nome di variabile> possono essere utilizzate per centrare una linea o stampare il valore di una variabile precedentemente definita o appartenente al questionario originale.

@SYSTEMDATE stampa il valore di SYSTEMDATE.

*[<valore di variabile1>, {<valore di variabile2>, <valore di variabile3>, ecc.}]
{:<Num. di cifre>}*

Le parentesi quadre che racchiudono un elenco di valori di variabili specificano che i contenuti della cella della tabella determinata dai valori saranno stampati nel punto in cui l'espressione tra parentesi appare.

Esempio

Il numero dei casi di epatite nella contea di Monroe tra le donne è

[MONROE, HEPATITIS, F]

Il numero dei casi di tutte le condizioni nella contea di Monroe è [MONROE,,].

Il numero totale dei casi di epatite è [,HEPATITIS,].

Il numero totale degli uomini è [, ,M].

Il totale generale è [, ,,]:3

{ <espressione matematica> };<Numero totale delle cifre>.<cifre decimali>

Le parentesi graffe indicano un calcolo che utilizza operazioni tra parentesi quadre o variabili o valori letterali inclusi tra parentesi graffe. Il numero totale di cifre, che include il punto decimale, e la posizione del punto decimale sono indicati alla destra delle parentesi graffe. Gli operatori permessi sono quelli elencati sotto "Operatori" nella sezione di riferimento di ANALYSIS.

Consultare il capitolo 24 del manuale elettronico per ulteriori esempi.

RETURN

Sintassi: RETURN

RETURN fa sì che il programma riprenda con il comando a partire dall' ultimo comando GOTO eseguito.

ROUTE

Sintassi: ROUTE <Nome del file>

ROUTE SCREEN

ROUTE PRINTER

Route apre il file specificato come file di destinazione dei risultati. Tutti i risultati ottenuti dal programma andranno nel file specificato, ad eccezione dei messaggi di errore selezionati. Due "file" speciali, "SCREEN" e "PRINTER", consentono di inviare i risultati, rispettivamente, allo schermo e alla stampante. Per cancellare un precedente ROUTE per un file o per una stampante, utilizzare il comando ROUTE SCREEN.

RUN

Sintassi: RUN <nome del file di programma>

RUN avvia un file di programma (.PGM) contenente dei comandi. Il programma leggerà i comandi dal file specificato fino a che non viene raggiunta la fine del file o viene incontrato un altro comando RUN. In questo caso il controllo del programma viene trasferito al secondo programma e non ritorna automaticamente al primo.

SAVE

Sintassi: SAVE <Nome di file>

SAVE immagazzina i comandi precedentemente inseriti in un file di programma per modificare o correggere la sequenza di comandi usando l'editing in EPED. Il numero di comandi salvati viene determinato dal valore di CMDSTACK, ed è usualmente 20. Può essere variato con il comando SET CMDSTACK = <un numero> fino a 100 righe di comando. SAVE salva i comandi del programma, non i dati, e i file .REC non vengono modificati.

SCATTER

Sintassi:

SCATTER <Variabile1> <Variabile2> {/R}{/X=<min>-<max> /Y=<min>-<max>}

SCATTER realizzerà un grafico a punti (scatter plot) con la Variabile1 sull'asse orizzontale e la Variabile2 sull'asse verticale. Se "/R" è incluso dopo i nomi di campo, una linea di regressione dei minimi quadrati verrà disegnata tra i punti relativi ai dati. I grafici a punti sono i più utili quando entrambi i campi contengono dati numerici continui come età o pressione arteriosa anziché informazioni codificate o del tipo Sì/No.

SELECT

Sintassi: SELECT <espressione>

SELECT vi consente di specificare un'espressione che deve essere vera per tutti i record da elaborare. Se la selezione attuale è "Età > 35", allora solo i record con età maggiore di 35 verranno elaborati. SELECT usato da solo senza alcuna espressione cancellerà tutti i precedenti comandi SELECT.

Le espressioni SELECT sono cumulative, tanto che le due espressioni:

```
SELECT AGE > 35
```

```
SELECT SEX = "F"
```

sono equivalenti a SELECT (AGE>35) AND (SEX = "F").

SELECT da solo elimina i precedenti comandi SELECT

Vedi pagina 64 o capitolo 11 del Manuale.

SET

Sintassi: SET {<Condizione = <Valore>}

SET vi permette di impostare i vari parametri di formattazione di ANALYSIS. SET da solo visualizza le configurazioni attualmente attive.

I comandi SET possono essere digitati dalla tastiera o posizionati nei file di programma. Sono utili specialmente quando posizionati in un file chiamato CONFIG.EPI, che è av-

viato automaticamente quando viene avviato ANALYSIS. Il posizionamento di comandi SET in questo file ha l'effetto di realizzare una configurazione permanente.

Le condizioni che influenzano la stampa ed i valori sono:

- **SET PRINTER** Si possono trasmettere i Grafici SET PRINTER in ANALYSIS alle stampanti compatibili a matrici di punti e alle stampanti laser ed ai plotter della Hewlett Packard compatibili. Con Epi Info, vengono forniti 30 driver di stampa. I nomi di questi file possono essere usati con SET PRINTER = <nome del driver>. I nomi dei driver iniziano con "\$" e terminano in .BGI. Per impostare ANALYSIS in modo che produca un file CGM, ad esempio, si introduce il comando, SET PRINTER = \$CGM. I driver forniti sono:

Nome File		Kbyte	Formato
\$BMP	BGI	17768	BMP (Windows 3.x) Format
\$CGM	BGI	11818	Computer Graphics Metafile
\$FX	BGI	20894	Epson FX80 Printer
\$PS	BGI	15518	Postscript
\$TIF	BGI	20862	Compressed TIF Format
\$UTIF	BGI	20862	Uncompressed TIF Format
\$WPG	BGI	9642	WordPerfect Graphics
\$CFX	BGI	21128	Epson 9-pin Color Printer
\$CLQ	BGI	23726	Epson 24-pin Color Printer
\$DJ	BGI	21792	Hewlett-Packard Deskjet
\$DJC	BGI	22252	Hewlett-Packard Color Deskjet
\$DXF	BGI	16084	AutoCad DXF Format
\$HP7090	BGI	16264	Hewlett-Packard Plotter 7090
\$HP7470	BGI	16232	HP Plotter 7470
\$HP7475	BGI	16336	HP Plotter 7475
\$HP7550	BGI	16336	HP Plotter 7550
\$HP7585	BGI	16368	HP Plotter 7585
\$HP7595	BGI	16368	HP Plotter 7595
\$IBMQ	BGI	20220	IBM Quietwriter Printer
\$IMG	BGI	20402	GEM IMG Format
\$LJ	BGI	23146	Hewlett-Packard Laserjet
\$LJ3R	BGI	23350	Hewlett-Packard Laserjet III
\$LQ	BGI	23492	Epson 24-Pin Printer
\$OKI92	BGI	18064	Okidata 92 Printer
\$PCX	BGI	17616	PCX (640 x 480) Format
\$PJ	BGI	24940	Hewlett Packard Paintjet
\$PP24	BGI	23562	IBM Proprinter, 24 Pin
\$TSH	BGI	22306	Toshiba 24-Pin Printer
\$VSHO	BGI	13112	ANSI NAPLPS/VideoShow Format

- **SET GRAPH.** Questo comando sceglie un driver per l'output grafico da trasmettere al file. Se scegliete SET GRAPH = \$CGM, ad esempio, il driver CGM verrà usato se (e soltanto se) il percorso (ROUTE) di trasmissione dell'output è verso il file. Tuttavia, quando l'output viene trasmesso alla stampante, verrà usato il driver specificato in SET PRINTER.
- **SET PORT** dirige il prodotto verso la porta corretta se la vostra stampante è collegata ad una porta che non sia quella parallela standard. I valori possibili sono LPT1, LPT2, COM1 e COM2.
- **SET BAUD** Imposta la frequenza di trasmissione alla porta. Ciò è necessario per i plotter e talora per le stampanti laser che usano una porta seriale. Le possibili scelte sono 300,600, 1200, 2400, 4800, 9600 e 19200.
- **SET PMODE** Le stampanti assistite da ANALYSIS per la stampa di grafici hanno una scelta di modi con cui controllare la dimensione dell'immagine e la risoluzione. I valori disponibili per differenti stampanti, come singoli caratteri, sono: per stampanti Epson ad aghi: 0=metà pagina; 1=pagina intera; 2=pagina intera in neretto. Le stampanti Epson LQ1500 possono funzionare nei modi precedenti ed anche: 3=Letter quality, mezza pagina; 4=Letter quality, pagine intera. Le stampanti Hewlett Packard LaserJet hanno: 0=metà pagina, bassa risoluzione; 1=metà pagina, media risoluzione; 2=metà pagina, alta risoluzione; 3=pagina intera, bassa risoluzione; 4=pagina intera, media risoluzione; 5=pagina intera, alta risoluzione. I plotter Hewlett Packard 7470A e 7475A hanno due modalità, che producono entrambi una pagina intera: 0=Draft quality; 1=Letter quality.
- **SET PAGE** controlla la dimensione della pagina in larghezza (numero di caratteri) ed in altezza (numero di righe). Normalmente questi parametri sono 23,80 per lo schermo e 66,80 per la stampante.
- **SET PAUSE** controlla l'interruzione della visualizzazione dei risultati dopo ogni pagina (ottimo per la stampante fogli singoli). La pausa è normalmente attiva per lo schermo e disattivata per la stampa.
- **SET PLINES** Rappresenta il numero di righe per pagina sul foglio usato dalla stampante. Di solito è necessario impostare la lunghezza della pagina per la stampante indipendentemente da quella dello schermo. Normalmente PLINES è 66, ma può essere impostata a 60 per molte stampanti laser.

Le condizioni che influenzano Tabelle, Frequenze o Elenchi sono:

- **SET CRITERIA** Quando CRITERIA = ON, i criteri dell'attuale selezione verranno stampati prima di ciascun elenco, frequenza o tabella. La configurazione normale è ON.
- **SET FILLCHAR** Controlla il carattere usato per riempire gli spazi vuoti presenti nei campi numerici scritti con il comando WRITE o LIST.
- **SET LINES** Quando LINES = ON, le tabelle verranno stampate con linee verticali ed orizzontali che separano le celle. La configurazione normale è OFF.
- **SET LISTREC** Quando LISTREC = OFF, i numeri dei record non saranno inclusi nel prodotto ottenuto con il comandi LIST. La configurazione normale è ON.
- **SET MAXCELLWIDTH** Questo è un numero che specifica la massima ampiezza delle celle della tabella per quanto riguarda il numero dei caratteri, di solito 15. Può

essere usato per limitare l'ampiezza delle celle, consentendo la visualizzazione di un'ampia tabella in una sola pagina. Se il numero specificato è troppo grande, verrà utilizzata l'ampiezza minima. Se MAXCELLWIDTH non viene specificato, l'ampiezza di ciascuna cella dipenderà dalla dimensione della variabile corrispondente.

- **SET PERCENTS** Quando PERCENTS = ON, le tabelle verranno stampate con la percentuale sotto ciascun conteggio. Vengono scritte prima le percentuali di riga, con una freccia indicante il denominatore sulla destra. Le percentuali in colonna sono stampate sotto le percentuali di riga. La configurazione normale è OFF.
- **SET STATISTICS** Quando STATISTICS = ON, verranno prodotte statistiche per ciascuna tabella specificata nel comando TABLES. Quando è OFF, risulterà solo la tabella. ON è l'impostazione standard.
- **SET SWITCH** Quando SWITCH = OFF, i valori "Y" ed "N" non saranno convertiti a "+" e "-". Normalmente ciò viene fatto in modo che la posizione di questi valori nelle tabelle sia conveniente, con i valori "+" sistemati per primi.

I comandi che influenzano l'elaborazione dei record o altre funzioni di ANALYSIS sono:

- **SET BREAKCMD** Il comando realizzato da BREAK, quando si manifesta una variazione nel campo(i) specificato nell'affermazione del comando, può essere impostato come nei seguenti esempi:
SET BREAKCMD = TYPE "Nuovo gruppo di età"
SET BREAKCMD = TYPE "\nAGE @AGE \n"
Se BREAKCMD non è specificamente impostato (SET), l'operazione realizzata da BREAK è di saltare alla pagina successiva durante la produzione di un elenco (LIST). Ciò capita solo quando i risultati vengono inviati (ROUTE) alla stampante o ad un file. Consultare BREAK per ulteriori informazioni.
- **SET BUFSIZE** Una porzione della memoria chiamata "buffer" viene usata per aumentare la velocità di lettura dei file. "Buffer" ampi consentono un'elaborazione più rapida, particolarmente dopo che un file è stato letto per la prima volta. I valori consentiti vanno da 0 a 32767.
- **SET CMDSTACK** Determina il numero di comandi inseriti dalla tastiera che possono essere richiamati con il tasto <freccia in alto> o salvati con il comando SAVE. Per esempio, SET CMDSTACK = 30 consentirà di trattenere gli ultimi 30 comandi inseriti. Normalmente CMDSTACK è impostato a 20, ma può essere configurato su di un numero minore, per conservare la memoria, o su di un numero maggiore se fate un uso estensivo delle caratteristiche di SAVE. Il numero massimo di comandi che possono essere salvati è 100.
- **SET IGNORE** Quando IGNORE = ON, i valori mancanti (missing) verranno ignorati nella costruzione delle tabelle. OFF è la condizione normale. Impostate IGNORE = ON per produrre tabelle 2 x 2 con variabili Si/No se vi mancano dei valori in alcuni record.
- **SET NOECHO** Quando NOECHO = ON, il prodotto di comandi quali TABLES e FREQ non sarà mostrato sullo schermo. La normale configurazione è OFF. NOECHO può anche essere usato come parametro dopo ciascun comando, come ad esempio in FREQ SEX /NOECHO.

- **SET PROCESS** ha tre configurazioni: DELETED, UNDELETED e BOTH. Esso controlla quali record verranno elaborati da ANALYSIS. Se il suo valore è UNDELETED, solo i record non cancellati verranno elaborati. Questa è la normale configurazione. Se è DELETED, solo i record cancellati verranno elaborati; ciò può essere utilizzato, ad esempio, per stampare una lista di tutti i record cancellati. SET PROCESS = BOTH determinerà l'elaborazione di tutti i record selezionati indipendentemente dalla cancellazione.

SORT

Sintassi: SORT {<nome della di variabile1> <nome della variabile2> ... <nome della variabile N>}

SORT vi consente di specificare la sequenza in cui verranno scritti i record con i comandi LIST e WRITE. Se non vi sono nomi di variabile specificati dopo il comando SORT, l'attuale operazione di classificazione (sort) viene annullata, e nei successivi prodotti i record appariranno nell'ordine che avevano nel file originale.

SUBSTRINGS

Porzioni di variabili a stringa, o di substrings, vengono indicate tra parentesi quadre nel seguente modo:

Nome della variabile [Posizione del primo carattere, Numero dei caratteri]

SUMFREQ E SUMTABLES

Sintassi: <Variabile da sommare> <Variabile alla quale si deve sommare> <Variabili aggiuntive alle quali si deve sommare>

SUMFREQ e SUMTABLES vengono utilizzati per l'elaborazione di record riassuntivi prodotti con il comando OUTPUT. Questi due comandi producono tabelle simili a quelle prodotte con i comandi FREQ e TABLES. FREQ e TABLES conteggiano il numero di record che soddisfano le condizioni specificate e pongono il valore in appropriate cellule delle tabelle. SUMFREQ e SUMTABLES aggiungono i contenuti della prima variabile specificata per tutte le combinazioni delle altre variabili e mettono il risultato in una o più tabelle.

TABLES

Sintassi: TABLES <Nome della variabile1> <Nome della variabile2> {<Nome della variabile3>...<Nome della variabile5>} {/SUM=<Nome della variabile>}

TABLES effettua una tabulazione incrociata delle variabili specificate e invia la tabella allo schermo, alla stampante, o ad un file. I valori della prima variabile appariranno sul margine sinistro della tabella, e quelli relativi alla seconda variabile, alla sommità della tabella. Normalmente le celle contengono il conteggio dei record corrispondenti alle etichette visualizzate sul margine. Se viene fornito il parametro /SUM, le celle rappresen-

tano le somme esattamente come nelle SUMTABLES. TABLES DISEASE COUNTY /SUM=COUNT dà gli stessi risultati di SUMTABLES COUNT DISEASE COUNTY.

Tutte le variabili che si trovano dopo le prime due elencate nel comando TABLES vengono utilizzate per la stratificazione. Per ciascuna possibile combinazione di valori di queste variabili, verrà prodotta una tabella separata (strato) per variabile 1 su variabile 2.

Il comando SET può essere usato per determinare il formato delle tabelle con l'omissione delle statistiche (SET STATISTICS = OFF), includendo i valori percentuali nelle tabelle (SET PERCENTS = ON), stampando linee verticali (SET LINES = ON), determinando la larghezza massima delle colonne (SET MAXCELLWIDTH = <numero>), omettendo la stampa dei criteri di selezione alla sommità di ciascuna tabella (SET CRITERIA = OFF) e prevenendo la conversione di "Y" e "N" delle variabili Sì/No a "+" e "-" (SET SWITCH = OFF). Consultate, per ulteriori dettagli, il comando SET.

TITLE

Sintassi: TITLE <Numero della riga 1-5> {"<testo del titolo>"}

TITLE imposta il titolo del testo specificato con il <Numero della riga> segnato. È facoltativo racchiudere il testo tra virgolette. L'utilizzo del numero delle righe da 1 a 5 consente la stampa di cinque righe di titoli alla sommità della tabella.

All'interno del testo del titolo, possono essere usati i seguenti comandi particolari:

\C	Il titolo verrà centrato.
\L	Il titolo sarà giustificato a sinistra.
\R	Il titolo sarà giustificato a destra.

@<Nome di variabile> Il valore attuale della variabile verrà inserito nel testo.

Un comando TITLE senza alcun testo pulisce quella riga ed anche ogni riga successiva.

TRIM

Sintassi: TRIM (<Nome di variabile>)

TRIM elimina gli spazi vuoti da un campo di testo. Ciò è utile nella combinazione di campi stringa per produrre etichette postali o nuove variabili per mostrarne un elenco compatto.

TYPE

Sintassi: TYPE "<Testo>" o TYPE <Nome di file>

Il comando TYPE possiede due significati, determinati dal fatto che l'informazione che segue sia compresa o meno tra virgolette. Se il testo viene fornito tra virgolette, TYPE invierà il testo allo schermo, alla stampante o ad un altro file di prodotto designato con l'ultimo comando ROUTE. Se il testo non è fra virgolette, Type presume che sia un nome di file e mostrerà a video il contenuto del file specificato.

UPDATE

Sintassi: UPDATE

Il comando UPDATE è identico al comando BROWSE, eccetto per il fatto che possono essere realizzati dei cambiamenti nel file di dati. Prima di utilizzare tale comando, assicuratevi di fare una copia del vostro file di dati, poiché le modifiche realizzate sono definitive! Questo è l'unico comando di ANALYSIS che realmente modifica il file originale di dati.

In UPDATE, sono disponibili le seguenti funzioni:

<Home><End> Questi tasti muovono al primo o all'ultimo record.

<F4> Full Screen. La pressione di tale tasto visualizza i record dell'insieme dei dati nel formato questionario come apparirebbero nel programma ENTER, o nel formato foglio elettronico come quando è attivato UPDATE. Il più familiare formato questionario fornisce un orientamento circa il significato delle domande. I tasti <F7> Prev e <F8> Next sono disponibili nel programma ENTER nel modo questionario.

<F5> Split Screen Questo comando mostra il formato foglio elettronico nella metà superiore dello schermo ed il formato questionario in quella inferiore. Una nuova pressione di <F5> ripristinerà il formato foglio elettronico a tutta pagina.

<F6> Delete. Questo comando segna un record come cancellato o riattiva un record cancellato, come già visto nel programma ENTER.

<F10> Done. Questo comando termina la sessione di UPDATE e ritorna al messaggio di richiesta di EPI>, pronto per il prossimo comando.

VARIABLES (<F3>)

Sintassi: VARIABLES

VARIABLES visualizza un elenco di nomi di variabili nell'attuale insieme di dati, con il tipo e la lunghezza di ciascuno, come pure i nomi ed i contenuti di ogni variabile definita dall'utente.

Il tasto funzione <F3> visualizza un elenco di variabili come un menù. Il cursore può essere spostato con i tasti freccia sulla scelta desiderata. Potete selezionare contemporaneamente numerose variabili con il tasto <+> sul lato destro della tastiera oppure rimuovere la selezione con il tasto <->. Quando viene premuto <Enter>, il(i) nome(i) della variabile verrà automaticamente inserito nel comando attuale.

WRITE

Sintassi: WRITE {RECFILE} {<nome di variabile> {" " <lunghezza della variabile>}} ecc.

WRITE invierà i record all'attuale file di prodotto nel formato da voi specificato. Potete specificare quali variabili verranno scritte, la lunghezza delle stesse, l'ordine in cui appariranno e se il file sarà o meno un file .REC. Possono essere specificati dei separatori od altro materiale da includere in ciascun record.

Se è specificato "RECFILE", viene creato un file Epi Info, contenente tutte le variabili specificate.

MISCELLANEA

? ? (IL COMANDO GET-INPUT)

Sintassi: ? <Messaggio per l'utente> ?

Questo è un comando interattivo di input. <Messaggio per l'utente> verrà visualizzato sullo schermo ed il sistema attenderà una risposta dalla tastiera. Qualsiasi cosa venga digitata dall'utente viene sostituita nella linea di comando in cui appaiono i due punti di domanda, come se facesse parte del comando originale.

Annotazioni

Capitolo 17

GLOSSARIO BREVE DEI TERMINI EPIDEMIOLOGICI FONDAMENTALI

ANALISI STRATIFICATA

Il programma ANALYSIS offre la possibilità di produrre una serie di tabelle stratificate, listando più di due variabili, con il comando TABLES. Le variabili poste dopo le prime due servono come base per dividere le tabelle in livelli o strati, uno per ogni combinazione della variabile posta dopo la seconda.

Data una serie di proporzioni che rappresentano un grado di esposizione crescente o decrescente (livello di dosaggio di un farmaco, n. di sigarette fumate, ecc.) e il numero di persone affette o non affette da una malattia in ciascun gruppo, si può testare l'odds (vedi) nei gruppi successivi comparandoli con il primo gruppo (in genere non esposti o "baseline"). È una tecnica utilizzata oltre che per studiare la relazione dose-risposta, per neutralizzare il ruolo di un eventuale fattore di confondimento. Per verificare se un'associazione è affetta da una distorsione e controllare gli effetti di una variabile potenzialmente confondente, è sufficiente ripartire i soggetti in strati, cioè in gruppi in cui l'esposizione al fattore di confondimento è identico. Suddividendo i dati in alcuni strati l'effetto confondente dovuto a fattori come, ad esempio l'età o il sesso, può essere annullato calcolando il rischio relativo a livelli di esposizione con equivalente fattore di confondimento: se esiste associazione si troverà un rischio relativo superiore a 1 (vedi) in ciascun strato. In STATCALC vengono calcolati gli odds ratios dei successivi livelli di esposizione, il chi quadrato per l'analisi stratificata di Mantel-Haenszel e il valore p relativo.

MISURE STATISTICHE DI DESCRIZIONE DI UNA DISTRIBUZIONE

Le misure comunemente usate per sintetizzare la tendenza centrale di una distribuzione sono la media, la mediana (il valore centrale delle osservazioni quando queste sono aggruppate in ordine ascendente), e la moda (il valore che occorre più di frequente). Le misure di dispersione della distribuzione includono il range e, per una distribuzione normale, la deviazione standard. Una distribuzione può essere descritta dai punti che la dividono in segmenti contenenti un uguale numero di osservazioni: questi sono i terzili, quartili, quintili, decili o percentili (il 50° percentile è la mediana). L'intervallo interquartile tra il quartile superiore e inferiore può essere usato come misura di dispersione.

ODDS RATIO

L'Odds può essere definito come il rapporto tra due probabilità: della probabilità che qualcosa si verifichi, sulla probabilità che non si verifichi.

Se in uno studio di coorte si osserva che 30 fumatori sviluppano una bronchite cronica e 20 no, l'odds per i fumatori in favore di sviluppare una bronchite cronica è di 30 a 20, o

60% a 40%, o 0.6 a 0.4, o 1.5 a uno, o - questa è il modo usuale espresso in epidemiologia - semplicemente 1.5. Questo è l'odds in favore della futura comparsa della malattia (odds della malattia). L'Odds può anche riferire del rapporto della probabilità che qualcosa è stato così nel passato, diviso la probabilità che non sia stato così. Se 40 persone con bronchite cronica sono stati fumatori e 10 no, l'odds nei pazienti in favore di essere stati fumatori è di 4 a 1 (odds dell'esposizione).

L'odds è usato nelle scommesse sui cavalli, per esempio 3 a 1 è l'odds a cui il bookmaker gioca contro la vittoria del cavallo, la probabilità che esso perda in rapporto alla probabilità che vinca. Si può interpretare anche come indice di "stranezza": quante volte è più "strano" che il cavallo vinca rispetto alla sconfitta.

L'odds è differente dalla probabilità: la probabilità che esca 2 con il lancio di un dado non truccato è di 1/6, l'odds che esca lo stesso numero è 1/5, cioè il rapporto tra il numero di casi vincenti contro quelli perdenti.

L'Odds Ratio (O.R.) è il rapporto fra un odds e un altro. È uno strumento largamente usato per valutare le associazioni tra variabili. L'O.R. rappresenta una approssimazione del rischio relativo solo se la malattia è rara nella popolazione generale e non si può calcolare il rischio relativo (studi caso-controllo), ma è sempre un indicatore del grado di associazione tra una malattia e l'esposizione in studio. Comparando l'odds di una malattia fra i fumatori con il corrispondente fra i non fumatori l'O.R. ci dice se la malattia è associata con il fumo e qual è il grado della forza di questa associazione.

Una formula utile per calcolare l'O.R. (senza calcolare prima separatamente l'odds della malattia e quello dell'esposizione) è ad/bc (tavole di contingenza, vedi), dove a rappresenta l'occorrenza combinata delle due variabili o categorie di cui si sta studiando l'associazione. Per questo l'odds ratio è anche chiamato "prodotto crociato".

REGRESSIONE

La regressione studia l'esistenza di una relazione di dipendenza tra una variabile e un'altra. I livelli della prima vengono stabiliti dal ricercatore (variabile indipendente), in corrispondenza viene misurata la seconda (variabile dipendente).

La tecnica della regressione lineare permette lo studio solo dell'associazione lineare tra due variabili e può essere rappresentata da una retta che attraversa la nuvola di punti individuati dalle due serie di misure.

RISCHIO

Si dice che il rischio è la probabilità di evenienza di un avvenimento (decesso, malattia), in un dato momento o durante un intervallo di tempo. Tra i rischi studiati in epidemiologia si può citare: la mortalità (l'avvenimento è il decesso), l'incidenza di una malattia o di un danno di salute ben definito. È evidente che questa probabilità non è la stessa per tutti gli individui di una popolazione. In effetti il rischio può variare in funzione di certe caratteristiche individuali, come il sesso, l'età, certi fattori socio-economici o certi parametri biologici: questi sono i fattori che si chiamano fattori di rischio. Sono fattori di rischio tutte le variabili legate statisticamente all'avvenimento studiato.

RISCHIO RELATIVO

Per giudicare se l'esposizione ad un fattore studiato modifica l'incidenza della malattia (il rischio), si utilizza un indice epidemiologico chiamato rischio relativo, che permette di comparare l'incidenza fra i gruppi di soggetti sottoposti a esposizioni differenti. Prendiamo ad esempio due gruppi di soggetti, uno che è stato esposto al fattore studiato, l'altro no. Il rischio relativo è il rapporto dell'incidenza nel gruppo esposto con l'incidenza nel gruppo non esposto.

$$RR = \frac{Ie+}{Ie-}$$

dove $Ie+$ è l'incidenza fra gli esposti ($a/a+b$ nelle tavole di contingenza, vedi)
 $Ie-$ è l'incidenza tra i non esposti ($c/c+d$).

Il valore del rischio relativo si interpreta nel seguente modo. Se il fattore studiato non gioca un ruolo causale, non deve esistere alcuna differenza di incidenza tra i soggetti esposti e i non esposti: in questo caso il rischio relativo sarà uguale a 1. Se è superiore a 1, significa che la presenza del fattore esercita un aumento della probabilità di insorgenza della malattia (o una diminuzione di questa probabilità se è inferiore a 1). Un rischio di 3 (o di 10) dovrà, dunque, essere interpretato così: i soggetti esposti al fattore di rischio hanno una probabilità 3 volte (o 10 volte) più elevata di avere la malattia che i non esposti. Il rischio relativo è dunque la misura del ruolo eziologico di un fattore di rischio.

SIGNIFICATIVITÀ STATISTICA

Si verifica la significatività statistica di una associazione per evitare di decidere se considerare i risultati dovuti al "caso". Molti test di significatività possono essere eseguiti per calcolare la probabilità (p vedi) che una associazione è tanto più forte di una osservata in un campione randomizzato (della dimensione attualmente in uso) selezionato da una popolazione nella quale questa associazione non esiste.

Il valore p è la probabilità di concludere che si tratti di una vera associazione, quando attualmente non lo è. I test di significatività non riguardano solo la variazione del campione randomizzato, ma anche gli errori aleatori di misurazione e altri processi casuali.

STUDIO CASO-CONTROLLO

“Uno studio caso-controllo è una inchiesta nella quale un gruppo di individui è selezionato in termini se hanno (casi) o non hanno (controlli) la malattia la cui eziologia viene studiata, i gruppi sono comparati per mezzo della esistenza presente o in passato delle caratteristiche giudicate di possibile rilevanza eziologica per la malattia.”

(B. MacMahon, T. Pugh, *Epidemiology principles and method*, 1970).

STUDIO DI COORTE

“Uno studio di coorte costituisce una importante forma di ricerca epidemiologica per testare ipotesi sulla causa delle malattie. Le caratteristiche distintive di uno studio di coorte sono:

1. Il gruppo o i gruppi di persone studiate (la coorte) è definita in termini di caratteristiche manifestate prima dell'insorgenza della malattia investigata.
2. Il gruppo in studio, così definito, è osservato lungo un periodo di tempo, in modo da determinare la frequenza della malattia entro il periodo di osservazione.”

(B. MacMahon, T. Pugh, *Epidemiology principles and method*, 1970).

TASSI

L'informazione sulla frequenza di un evento o attributo, in un gruppo o in una popolazione, è comunemente riassunta calcolando un tasso. Il numero di eventi, o il numero di individui con una caratteristica (il numeratore del tasso) è diviso con un adatto denominatore (es. il numero di soggetti di una popolazione) e moltiplicato per 100, 1000 o altra costante conveniente. Ciò permette di controllare l'effetto della dimensione della popolazione sulla frequenza dell'evento o attributo. Il tasso di incidenza misura la frequenza di eventi occorsi in uno specifico periodo di tempo. Il tasso di mortalità è un tasso di incidenza che misura la frequenza dei decessi.

Tutti i tassi sono rapporti, calcolati dividendo un numeratore con un denominatore. Alcuni tassi sono proporzioni: cioè il numeratore è contenuto nel denominatore. Il corretto uso del termine tasso è sfortunatamente controverso. Alcuni autori attribuiscono l'uso del termine ai rapporti che riflettono un cambiamento relativo fra due quantità, altri lo restringono solo ai rapporti che rappresentano un cambiamento nel tempo; secondo questa teoria, un tasso di prevalenza non è un vero tasso.

TASSO DI ATTACCO

In certe circostanze una popolazione può essere a rischio di malattia solamente per un limitato periodo di tempo. Ciò può essere dovuto alla breve durata nel tempo in cui opera il fattore eziologico o dalla breve durata di una epidemia, o che il rischio è limitato in una fascia di età particolare. In queste circostanze si esprime l'incidenza, limitata ad un breve lasso di tempo, con il nome di tasso di attacco senza l'indicazione del periodo.

TAVOLA A DOPPIA ENTRATA (O TAVOLA DI CONTINGENZA)

Per illustrare sinteticamente la distribuzione dei soggetti rispetto la presenza o l'assenza di due (o più) caratteristiche (fattore di esposizione e evento dannoso), si adopera la tavola di contingenza in cui ciascun individuo viene collocato in una cella. Le espressioni nella tavola possono essere frequenze (numero di individui), percentuali o altre proporzioni, tassi o medie, la misura espressa in ciascuna cella viene indicata con una lettera a, b, c, e d.

Evento

<i>Fattore</i>	<i>Presente</i>	<i>Assente</i>	<i>Totale</i>
Presente	a	b	a+b
Assente	c	d	c+d
Totale	a+c	b+d	a+b+c+d

Notate che l'interpretazione del rischio relativo dipende dall'orientamento della tabella e che non tutti i rischi relativi calcolati saranno validi. Il rischio relativo risulta dal rapporto tra la proporzione del primo fattore della prima linea comparato con la proporzione del primo fattore della seconda linea $[(a/a+b)/(c/c+d)]$. Se l'orientamento dei fattori non rispetta la sequenza indicata (TABLES variabile esposizione variabile evento), l'indicatore risulta invalidato. Contrariamente l'Odds Ratio, indipendentemente dall'orientamento degli elementi della tavola, fornisce sempre lo stesso risultato (ad/bc).

TEST CHI QUADRATO

Il test Chi quadrato è utilizzato per analizzare i dati disposti in categorie discrete (variabili qualitative, vedi) e quando le frequenze attese sono sufficientemente grandi (non più del 20 % dei valori attesi nelle varie celle deve essere inferiore a 5). Il numero delle classi può essere 2 o più. La tecnica permette di esaminare la "bontà della conformità", per stabilire se esiste una differenza significativa tra numero osservato di individui (o misurazioni) che cadono nelle varie categorie, e numero atteso di individui (o misurazioni) basato sull'ipotesi nulla: che non vi è diversità di frequenza degli individui (o misurazioni) fra le varie classi. ANALYSIS mette a disposizione tre valori diversi di Chi quadrato: la formula non corretta, la correzione di Yates e di Mantel-Haenszel, con i relativi valori p.

VALORE-P

La probabilità p è chiamata anche grado di significatività statistica (vedi).

Il suo calcolo permette di precisare il rischio potenziale di errore che accompagna il rigetto dell'ipotesi nulla: se p vale 0,01 (o 0,001) si dice che la differenza è significativa all'1% (o 1 per mille). È il famoso p che si trova nelle pubblicazioni scientifiche: è più rassicurante dire che esiste una differenza significativa se si può precisare che il rischio d'errore di una tale affermazione è inferiore a 1 su 1000, che dire che si hanno 5 probabilità su 100 di sbagliarsi.

VARIABILI

Certi fenomeni non possono essere compresi che in termini di presenza o assenza: si ha una malattia o non si ha, si è vivi o morti, si è ricercato o no. Le variabili che misurano questo tipo di fenomeni sono chiamate variabili qualitative. I valori suscettibili di essere compresi nella variabile costituiscono le modalità di quella variabile (un individuo può

essere celibe, sposato, vedovo o separato per esempio: la variabile "stato civile" ha così quattro modalità).

Altri fenomeni si esprimono sotto forma di una grandezza più o meno elevata, che si misura per mezzo di una scala divisa in unità: l'altezza di una persona (in metri o centimetri), la concentrazione di glicemia (in grammi o milliequivalenti). Tali variabili sono chiamate quantitative.

Una variabile misurata sotto forma quantitativa può sempre essere ulteriormente trasformata in variabile qualitativa, raggruppando in classi i valori che essa prevede (vedi "Ricodificare i dati" a pag. 68). Si può misurare l'altezza di un soggetto in centimetri e classificare in seguito i soggetti in classi di piccola, media e grande statura, fissando le relative soglie. Tali variabili vengono chiamate semi-quantitative (o ordinali in quanto esiste una relazione di ordine tra le differenti modalità).

INDICE DEGLI ARGOMENTI

Note all'edizione italiana

INTRODUZIONE AL CORSO

CHE COS'É EPI INFO
OBIETTIVO GENERALE DEL CORSO
OBIETTIVI SPECIFICI
Abilità di I livello
Abilità di II livello
Abilità di III livello

I PROGRAMMI DI EPI INFO 6

INSTALLAZIONE E AVVIO DI EPI INFO

INSTALLAZIONE
MESSA A PUNTO DELLA STAMPANTE
AVVIO DI EPI INFO
EPI INFO DA WINDOWS
CREAZIONE DEL GRUPPO: "PROGRAMMI DI EPI INFO" IN WINDOWS
VISUALIZZARE E MODIFICARE FILE DAL MENU PRINCIPALE DI EPI6

CREARE QUESTIONARI USANDO L'EDITORE EPED

INTRODUZIONE
IMPOSTARE EPED PER FARE UN QUESTIONARIO
INDICARE LA POSIZIONE PER L'INSERIMENTO DEI DATI NEL QUESTIONARIO
TIPI DI CAMPI IN EPI INFO
AIUTARE EPI INFO A CREARE NOMI DI VARIABILI SIGNIFICATIVI
LABORATORIO: EPED
ESEMPIO: andiamo a creare un questionario che intitoliamo:

COSA BISOGNA SAPERE SULLE BASI DI DATI

LA DEFINIZIONE DI UNA BASE DI DATI E LE SUE COMPONENTI FONDAMENTALI
LE TABELLE
ORGANIZZAZIONE
RELAZIONI FRA DUE TABELLE
RELAZIONI FRA PIÙ TABELLE
INTERROGAZIONI
SELEZIONI MEDIANTE OPERATORI LOGICI
CALCOLI
PRESENTAZIONI
VISUALIZZAZIONI
SCHEDE
REPORT
ETICHETTE

INSERIRE I DATI USANDO IL PROGRAMMA ENTER

INTRODUZIONE
AVVIARE ENTER
INSERIRE O MODIFICARE I DATI
CREARE UN NUOVO FILE DI DATI (.REC)

L'INSERIMENTO DEI DATI
SOMMARIO DEI TASTI FUNZIONE
MUOVERSI FRA UN RECORD E L'ALTRO
CERCARE RECORD RISPONDENTI AD ALCUNI REQUISITI (MATCHING CRITERIA)
MODIFICARE I RECORD GIÀ INSERITI
CANCELLARE E RECUPERARE I RECORD CANCELLATI
RISTRUTTURARE UN FILE DI DATI
STAMPARE IL QUESTIONARIO E I DATI
LABORATORIO: ENTER
Inserire i dati in CORSO.REC
Ristrutturare CORSO.REC modificando il questionario
Esempio di navigazione

IL PROGRAMMA CHECK: CONTROLLO

AVVIARE CHECK
MINIMI E MASSIMI (MIN/MAX) <F1> <F2>
Esempi
VALORI PERMESSI (LEGAL) <F6>
INSERIMENTO OBBLIGATORIO (MUST ENTER) <F4>
RIPETIZIONE (REPEAT) <F3>
CONDIZIONI PER SALTARE ALCUNI CAMPI (SKIP PATTERNS) <F7>
CAMPI COLLEGATI E CODIFICA AUTOMATICA <F5> E <F8>
COMANDO PER LA MODIFICA DEL CAMPO <F9>
CAMBIARE I NOMI DELLE VARIABILI
INTRODUZIONE DI COMMENTI AI "VALORI PERMESSI"
RICERCA AUTOMATICA
I FILE DI AIUTO PER I CAMPI DI ENTER
LABORATORIO: CHECK
Collegamento di due campi
Valori permessi (legal)

ANALYSIS: LISTE, FREQUENZE, TABELLE, STATISTICHE E GRAFICI

INTRODUZIONE
PER CONOSCERE ANALYSIS
Lezioni
IL SISTEMA DELLE GUIDE (HELP)
IL PRIMO PASSO: LEGGERE UN FILE
ELENCCARE I RECORD
FREQUENZE
TAVOLE DI CONTINGENZA
ANALISI STRATIFICATE
IL COMANDO MEANS (MEDIE)
Nota sui valori mancanti
TITOLI
GRAFICI
STAMPARE I RISULTATI O TRASFERIRLI IN UN FILE
RICHIAMARE I COMANDI DI ANALYSIS
LABORATORIO: ANALIZZIAMO IL FILE ORIENTE.REC

SCRIVERE PROGRAMMI E IMPOSTARE I DATI PER ANALYSIS

INTRODUZIONE
SCRIVERE PROGRAMMI
SALVARE E RICHIAMARE I COMANDI DI ANALYSIS
ACCESSO AL DOS DIRETTAMENTE DA ANALYSIS
ORDINARE I RECORD: IL COMANDO SORT

SELEZIONARE I RECORD DA ANALIZZARE
ESAMINARE E MODIFICARE I RECORD CON BROWSE E UPDATE DI ANALYSIS
 BROWSE (sfogliare)
 UPDATE (aggiornare)
DEFINIRE NUOVE VARIABILI
COME SONO ESEGUITI I COMANDI
ASSEGNARE UN VALORE SINGOLO: IL COMANDO LET
RICODIFICARE I DATI
PORRE CONDIZIONI CON IF/THEN
DATE
LABORATORIO: SCRIVIAMO UN PROGRAMMA

ANCORA SU ANALYSIS

MODIFICARE ALCUNE IMPOSTAZIONI DI ANALYSIS
I COMANDI SET PER LA STAMPANTE E I GRAFICI
CREARE UN NUOVO FILE EPI INFO
RISTRUTTURARE I RECORD
I COMANDI ? ?, TYPE, ECHO, BREACK, IMMEDIATE
RECORD RIASSUNTIVI: OUTPUT, SUMFREQ E SUMTABLES
GRAFICI DI RECORD RIASSUNTIVI
UNIRE I DATI IN CAMPI MULTIPLI NELLO STESSO FILE: IL COMANDO COMBINE
CONTROLLO DEL PROGRAMMA: I COMANDI RUN, GOTO E RETURN
LABORATORIO: IL PROGRAMMA ORIENTE.PGM

STATCALC ED EPITABLE: CALCOLATORI STATISTICI

INTRODUZIONE
AVVIO DI STATCALC
ANALISI DI TABELLE SINGOLE E STRATIFICATE (2 x 2 FINO A 2 x 10)
DIMENSIONE DEL CAMPIONE E POTENZA
CHI-QUADRATO PER TREND
STAMPA O INVIO DEI RISULTATI AD UN FILE
EPITABLE

VALIDATE: DOPPIA IMMISSIONE DI DATI E VERIFICA

INTRODUZIONE
COME AVVIARE VALIDATE
USARE ENTER PER LA CONVALIDA DEI DATI

EXPORT: PRODURRE FILE PER ALTRI SISTEMI

INTRODUZIONE
 Tipi di file prodotti da EXPORT
UTILIZZO DI EXPORT
LANCIO DI EXPORT DA DOS
LABORATORIO: ESPORTIAMO UN FILE

IMPORTARE DATI DA ALTRI PROGRAMMI

INTRODUZIONE
AVVIO DI IMPORT DA DOS
LABORATORIO: IMPORTIAMO UN FILE DI UN ALTRO PROGRAMMA

EPIGLUE: UN SISTEMA INFORMATIVO EPIDEMIOLOGICO E DI MENU

INTRODUZIONE
L'USO DELL' EDITORE TESTI INCORPORATO IN EPI6
LA COSTRUZIONE DEL MENU PERSONALIZZATO
I FILE DI IMPOSTAZIONE DEI MENU (.MNU)

IMPOSTAZIONE DEI MENU E DELLE VOCI DEI MENU

MENUIITEM
MENUIITEM SEPARATOR
MENUEDIT
POPUP

LA TABELLA DI STRINGHE

STRINGTABLE

I COMANDI DI BLOC

FILE GUIDA IPERTESTUA

LABORATORIO: CURIOSIAMO IN EPIGL

SINTESI DEI COMANDI DI EPI IN

COMANDI DI EP

Tasti funzionali <ctr
Tasto <scrl loc
<F1> il comando HELP (aiut
<F2> i comandi FI
<F3> il comando EPIA
<F4> i comandi di tes
<F7> i comandi di ricer
<F8> il blk o comandi di bloc
<F9> il comando SA
<F10> il comando DO
Altri comandi a tastie

COMANDI DI ANALYS

BAR
BREAK
BROWSE
CLOSE
CLS
CMD
COMBINE
COMMANDS
COPY
DEFINE
DOS
ECHO
ERASE
FOOTER
FORM
FREQ
<F10>
GENERATE
GOTO
HEADER
HISTOGRAM
IF
IMMEDIATE
LET
LINE
LIST
LOAD
MATCH
MATCHED
MEANS
NEWPAGE
OPERATORI
OUTPUT
PIE
PROCESS
QUIT
READ

RECODE
REGRESS
RELATE
RENAME
REPORT
RETURN
ROUTE
RUN
SAVE
SCATTER
SELECT
SET
SORT
SUBSTRINGS
SUMFREQ e SUMTABLES
TABLES
TITLE
TRIM
TYPE
UPDATE
VARIABLES
WRITE
MISCELLANEA
?? (il comando Get-Input)

GLOSSARIO BREVE DEI TERMINI EPIDEMIOLOGICI FONDAMENTA

ANALISI STRATIFICATA
MISURE STATISTICHE DI DESCRIZIONE DI UNA DISTRIBUZI
ODDS RA
REGRESSI
RISCHIO
RISCHIO RELATIVO
SIGNIFICATIVITÀ STATISTICA
STUDIO CASO-CONTROLLO
STUDIO DI COORTE
TASSI
TASSO DI ATTACCO
TAVOLA A DOPPIA ENTRATA (O TAVOLA DI CONTINGENZA)
TEST CHI QUADRATO
VALORE-P
VARIABILI.

