

MANUALE
DI
STATISTICA TEORICA
E
DEMOGRAFIA

AD USO DEGLI ISTITUTI TECNICI

DI

NAPOLEONE COLAJANNI

PROFESSORE ORDINARIO DI STATISTICA
NELLA R. UNIVERSITÀ DI NAPOLI

Edizione Quarta con due appendici

- I. *Statistiche economiche, intellettuali, politiche e morali*
di N. COLAJANNI.
- II. *Brevi cenni di qualche metodo di elaborazione dei dati*
statistici di ALFREDO NICEFORO.



NAPOLI
LUIGI PIERRO, EDITORE
Piazza Dante, 76
1920

Tutti i diritti di traduzione, riduzione e riproduzione sono riservati.

Gli esemplari non muniti dalla seguente firma sono falsificati:

C. Colajanni

PARTE PRIMA
STATISTICA TEORICA

LIBRO I.

Generalità.

CAP. I.

STORIA DELLA STATISTICA *.

1. I. *I fatti*. — Tra le discipline di origine recente poche ve ne sono, che hanno preso uno sviluppo tanto considerevole quanto la statistica.

Lo sviluppo è stato, per così dire, estensivo ed intensivo; dappoiché se la statistica penetra e s'infiltra negli studi di ogni genere, mercè sua gli argomenti più svariati vengono al giorno d'oggi trattati colla maggiore diligenza possibile; analizzati nei più minuti dettagli.

La statistica, però, si considera come una disciplina di data recente soltanto sotto l'aspetto della formulazione delle teorie; poichè in quanto ai *fatti*, agli elementi suoi costitutivi, essa rimonta alla più remota antichità. Invero il *fatto*, l'azione istintiva si trova alle origini della civiltà; ma solo dopo lunghi secoli, dei fatti si indagano i rapporti causali e l'empirismo consuetudinario si muta in un sapere organizzato, in un processo razionale.

Non si esagera quando si afferma che sin da quando vi furono Stati, che raggiunsero la fase della *scrittura* e dell'*aritmetica* vi furono numerazioni di cittadini e cataloghi e ap-

* Ho cercato di ridurre e di semplificare quanto più mi è stato possibile questo *Manuale* ad uso degli Istituti Tecnici; perciò ho soppresso quasi tutte le note e le citazioni, Per la bibliografia rinvio gli studiosi al *Manuale di Statistica teorica* ed al *Manuale di demografia* (due volumi di oltre 1100 pagine).

prezzamenti dei loro averi, affinché si potesse ordinare gli eserciti e i tributi. Sin dai primi momenti della vita politica della collettività umana, quindi, si fecero delle operazioni che adesso denominiamo *statistiche*, senza che vi fosse il più lontano accenno alle teorie ed al metodo della *Statistica*.

Le prime rilevazioni dei fatti e dati statici ci vengono specialmente dai censimenti e dai catasti; dei quali si trova fatta menzione nei libri della più remota antichità: nella Cina, nelle Indie, nell'Assiria, nell'Egitto, nella Grecia, in Roma.

Il Censimento in Roma, con particolarità, venne fatto più regolarmente: ed anche ivi sotto l'Impero si riscontrano funzionari stabili, che avevano incombenze statistiche e istituzioni, che hanno già qualche rassomiglianza con quelle, che assai più tardi si dovranno svolgere fra i popoli più civili.

Nella notte del medio evo le tracce delle raccolte di fatti d'indole statistica assunsero l'impronta del carattere generale dell'epoca nelle *diptyca*, cioè nelle liste che il clero cominciò a mantenere nelle Chiese e nelle quali si notavano le morti, le nascite, i matrimoni.

A misura che il Medio evo si dilegua, i fatti d'indole statistica pigliano il sopravvento. Epperò non è per accidente se li riscontriamo in onore crescente nelle repubbliche italiane, a Milano, a Firenze, a Bologna, ecc., che rappresentarono allora una somma considerevole di benessere economico, di sviluppo intellettuale.

Merita specialissima menzione, però, Venezia. Nel secolo XII la repubblica delle Lagune ordinò il catasto; nel secolo XIII (1268 e 1296) ordinò che ogni residente o ambasciatore alle Corti dei vari Stati dovesse disporre e fare un ragguaglio dello stato fisico e sociale del paese donde veniva: sono le famose *relazioni*, che hanno preceduto nella finalità ed anche nel metodo, i rapporti, che i consoli d'Inghilterra e degli Stati Uniti pei primi e poi quelli di tutte le altre nazioni oggi rivolgono ai Governi per informarli di tutto ciò che si riferisce alle condizioni economiche degli Stati e delle Città presso i quali sono accreditati. Venezia di buonora ordinò le anagrafi nel secolo XIV ed era il magistrato della Sanità, che raccoglieva i dati sulla mortalità. Ma ciò che la distingue soprattutto sono i suoi *censimenti*, il primo dei quali rimonta all'anno 1388, e che si continuarono per lunga serie di anni.

E censimenti, catasti, fatti vari d'indole statistica si riscontrano sino alla fine del secolo XVIII in America, nel Portogallo, nella Svezia, in Russia, in Prussia, nell'Impero Austriaco, ecc.

Col 1790 si ha il primo censimento decennale nella Repubblica degli Stati Uniti, come prescrive la Costituzione. Con questo avvenimento diamo termine a questa semplice rassegna di *fatti*.

La teoria è già sorta da un pezzo e si va elaborando; e *fatti*, *teorie* ed *istituzioni* s'intrecciano ed agiscono e reagiscono reciprocamente gli uni sulle altre, promovendo lo sviluppo e la sistemazione del metodo e di quella parte della Statistica applicata, che si considerò come scienza della *Statistica*.

2. II. *Le teorie*. — La storia delle teorie è assai difficile perchè le scoperte si concatenano sempre in modo complicatissimo e si trovano accenni vaghi su di un argomento in uno scrittore, su di un altro in un altro e così via di seguito, senza che si possano in generale stabilire delle divisioni nette fra i differenti periodi, tra le diverse teorie, tra i vari scrittori. Ciò che uno accenna, un altro sviluppa o modifica; e più tardi qualcuno riprende, quasi *ad pristinum*.

A voler fare la storia completa e nell'ordine cronologico dei vari scrittori che ne hanno trattato occorrerebbero parecchi volumi ed un intero corso di lezioni. Nè a tutti gli scrittori, nè a tutte le nazioni riesce facile assegnare il posto, che occupano in quanto a benemeranza.

Il nome stesso di *Statistica* e la sua etimologia hanno dato luogo a controversie.

Statistica, dice Messedaglia, viene da *Status* nel senso di *respublica*, stato politico e società civile. Donde: *statista*, che tra noi vuol dire uomo di Stato, o studioso di cose di Stato.

Tale è l'etimologia vera e storica.

Secondo un'altra etimologia, il nome della nostra disciplina deriverebbe da *Status*, Stato, nel senso di *conditio rerum*, situazione e condizione attuale di cose.

Del resto si sa che il nome dato ad una disciplina alle origini talora non corrisponde più al suo contenuto per la evoluzione successiva, ch'essa ha subito.

Tralasciando di far menzione dei numerosi scrittori che cooperarono allo sviluppo della teoria della statistica, qui mi limito a ricordare soltanto i nomi di Achenwal, di Sussmilch e di Quetelet. Da rilevare le affinità e le differenze tra i due ultimi, veramente sommi. Entrambi videro la *regolarità* nello svolgimento de' fenomeni demografici o sociali. Ma Sussmilch tale *regolarità* attribuì all'ordine divino, che a tutte le cose presiede; Quetelet, invece, l'attribuì alla forza stessa della natura. Concetto che viene indicato dal titolo stesso di una delle sue maggiori opere: *La fisica sociale*.

3. III. *Le istituzioni*. — Man mano che i privati e le pub-

bliche amministrazioni hanno riconosciuto la utilità immensa di avere notizie esatte e precise sulle condizioni della società, e su quelle dello Stato ch'è il suo organo, si è sentito il bisogno di procedere regolarmente alla raccolta e alla elaborazione delle notizie e de' dati statistici. Così lo Stato e i Comuni hanno proceduto alla sistemazione degli uffici di statistica e alle loro funzioni. Lo Stato non solo ha proceduto alla organizzazione degli uffici, ma si è incaricato pure della pubblicazione dei materiali raccolti ed elaborati.

Queste operazioni sono divenute periodiche e regolari, ma non sono mancate altre straordinarie, quali sono le inchieste.

I privati e le associazioni hanno variamente contribuito all'opera dello Stato, e si deve ad essi principalmente se si è riusciti ad organizzare i congressi e le pubblicazioni internazionali; nonchè a far sorgere un *Istituto internazionale*. Il quale però potrà assurgere a tutta la sua importanza solamente quando interverrà lo Stato; e allora darà risultati che sinora non ha potuto dare non ostante le grandi benemeritenze dei suoi promotori ed organizzatori.

Lo Stato aveva l'obbligo d'intervenire e di organizzare esso stesso la statistica, perchè ai privati e alle associazioni per quanto volenterosi e potenti non sarebbe mai riuscito di procedere con uniformità e regolarità a certe indagini e alla raccolta di certe notizie. Lo Stato solo ha in mano taluni elementi: *importazioni, esportazioni, registri giudiziari, registri dello stato civile* ecc.; lo Stato solo, inoltre, ha il diritto di farsi dare le notizie ch'è necessario raccogliere tra i privati.

Appena le ricerche statistiche divengono funzioni dello Stato, sorgono gli *organ*i relativi.

Gli uffici e le direzioni generali della statistica si sono fondati e sviluppati gradatamente, più o meno rapidamente e completamente in Europa e altrove dalla fine del secolo XVIII ai nostri giorni.

4. *Sviluppo delle istituzioni*.—Nella storia di questo sviluppo Mayr distingue tre periodi.

Nel *primo periodo*, che abbraccia all'incirca la seconda metà del secolo XVIII ed i primi anni del XIX, il bisogno di documenti statistici ufficiali comincia a farsi sentire. L'arte della rilevazione dei dati è però ancora affatto embrionale. Al centro della amministrazione dello Stato si designano, è vero, i singoli argomenti sui quali si vogliono ricevere delle dichiarazioni, ma appena si sa rendersi ragione dei metodi di rilevazione. Si esclude come principio di massima ogni pubblicazione di statistiche ufficiali o la si permette soltanto

con speciali cautele. Le tavole statistiche fanno parte dei segreti di Stato.

Il *secondo periodo* si può far decorrere dalle istituzioni di speciali *uffici di statistica*. Il materiale statistico, che si è raccolto nel corso del primo periodo, non è più in grado di esser dominato dalle autorità centrali esistenti, specialmente da quelle dell'amministrazione interna. È allora che si fondano gli uffici di statistica. In sulle prime essi hanno una posizione molto limitata. Di regola anche in questo periodo rimane alle autorità locali la facoltà di raccogliere ed anche di elaborare coi processi tecnici della statistica i fatti osservati. Gli uffici statistici non assumono in sostanza che l'ultimo concentramento e la revisione del materiale già disposto in tavole dalle autorità secondarie, e procurano un po' per volta la pubblicazione del medesimo. Dipende poi assolutamente dalla persona, che regge l'ufficio di statistica, il sapere contemporaneamente a questa pubblicità, collegarvi, in quanto ciò sia possibile, un'ulteriore elaborazione, in pro' della pratica e della scienza, dei risultati ottenuti.

Nel *terzo periodo* lo sviluppo della statistica ufficiale non è soltanto estensivo; cioè non vengono soltanto aperti alla statistica nuovi campi finora rimasti inesplorati dalla osservazione in massa quantitativa; ma è soprattutto intensivo e rappresentato dal lento e graduale costituirsi dell'arte statistica. Allora soltanto si compie la separazione già precedentemente descritta, fra i rilievi statistici da un lato e la elaborazione dall'altra, per così dire, della materia grezza, che se n'è ricavata. Quest'ultima operazione ormai concentrata negli uffici di statistica, raggiunge un perfezionamento inaspettato e impreveduto.

Si comprende che tra questi periodi non ci sono tagli netti e che nei diversi paesi d'Europa c'è una grande diversità di sviluppo; in guisa che mentre l'uno è, ad esempio nel 1° periodo, un altro si trova nel 2° o nel 3°. Tutto questo non riguarda che la statistica ufficiale; ma si capisce che tra non guari un *quarto* periodo dovrà svilupparsi ed è quello in preparazione, ad iniziativa sinora dei privati e delle associazioni, e sarà quello in cui la *statistica internazionale*—sinora opera dei Congressi, dello *Istituto* e delle relative pubblicazioni—diverrà anch'essa ufficiale. L'ultimo Congresso di Vienna (1913) lo fa sperare.

5. *Internazionalizzazione, specializzazione, e localizzazione della statistica*.—L'*internazionalizzazione* della statistica è una spiccata tendenza del momento presente, che ha bisogno di

esplicarsi maggiormente, affinchè si riesca alla maggiore unità possibile tra tutti gli Stati civili nella rilevazione, elaborazione ed esposizione dei dati.

In conformità di tale tendenza i cultori della statistica hanno sentito la immensa utilità d'intendersi tra loro; perciò hanno di buon'ora organizzato i *Congressi*, che preparano la internazionalizzazione degli *uffici della statistica* ed agevolano le comparazioni internazionali, che hanno specialissima importanza.

L'idea di un congresso internazionale venne suggerita a Quetelet dalla prima Esposizione Universale di Londra nel 1851; e il primo congresso si riunì a Bruxelles nel 1853. Nel 1885 sorse l'*Istituto internazionale di statistica* a Londra e come segretario generale venne scelto il Bodio, sotto la cui direzione si pubblicò l'importantissimo *Bulletin de l'Institut international de statistique*. Ora il segretario generale è l'olandese Verijn-Stuart.

A questo processo d'internazionalizzazione procede parallela la più intensa specializzazione e divisione del lavoro nella organizzazione interna degli uffici di statistica; e qualche ramo ha già preso tale importanza, che si è reso autonomo.

Così mentre da pertutto c'è già una *Direzione generale della statistica* alla dipendenza di questo o di quell'altro Ministero—in Italia alla dipendenza del Ministro di Agricoltura—come un esponente della importanza che ha acquistato la questione sociale e come una specializzazione degli *Uffici generali di statistica* sorgono gli *Uffici del lavoro* e si sviluppano attorno ad essi altre istituzioni, come il *Consiglio superiore del lavoro* in Francia, nel Belgio ed ora anche in Italia.

Il primo di questi uffici sorse nel Massachusetts nel 1869 e ne assunse la direzione Carrol D. Wright.

L'utilità di questi *Uffici del lavoro* è grande sia per la legislazione sociale; sia per la ricerca delle condizioni del lavoro, per i salari, ecc. perciò successivamente l'Inghilterra la Francia, il Belgio, l'Austria ecc., ed ultima l'Italia in seguito alla mia iniziativa parlamentare, li hanno imitati.

Il processo estensivo del quarto periodo d'*internazionalizzazione* ha condotto alla istituzione di un *Ufficio internazionale del lavoro* a Berna, che non ha, però, carattere ufficiale e del quale fanno parte uomini di diverso colore politico ma che s'interessano vivamente ai problemi del lavoro. Dal 1902 pubblica un *Bullettin de l'office international du travail*. Poste e telegrafi hanno pure a Berna un *Ufficio internazionale*

ufficiale. L'importanza somma di questo processo d'*internazionalizzazione* si potrà meglio apprezzare ricordando che molte riforme in favore dei lavoratori non sono possibili, perchè non essendo adottate dagli altri Stati pongono gli industriali delle nazioni, che le mettono in pratica, in condizioni d' inferiorità nella concorrenza sul mercato mondiale. D' onde il relativo Congresso internazionale di Berlino nel 1891 promosso da Bismark. Della internazionalizzazione delle leggi sul lavoro si occupò anche la Conferenza per la pace di Parigi (1919).

Parallelo al movimento della *internazionalizzazione* degli *Uffici di statistica* e della loro *specializzazione* è quello della loro *localizzazione*. Così oggi sorgono dappertutto *Uffici del lavoro* presso i grandi Municipi; ma non mi sembrano destinati tutti a fiorire. In Italia sembrano addirittura un prodotto della moda e dello spirito d'imitazione. I vari uffici di Statistica dei Municipi, però, possono rendere e rendono dei grandi servizi alla scienza ed alla società. Essi, infatti, hanno a loro disposizione un grande materiale di prima mano, che raccolgono direttamente e continuamente e che possono elaborare e pubblicare molto tempo prima degli Uffici Nazionali.

6. *Organizzazione della Direzione generale della Statistica in Italia*. — E' doveroso dare qualche cenno sulle istituzioni ufficiali di Statistica in Italia. Un Ufficio di Statistica troviamo in Sicilia fin dal 1832, che aveva come organo il *Giornale di Statistica di Palermo*, in cui si svolsero discussioni interessanti e cui collaborarono scrittori della fama di Ferrara, degli Amari, ecc. Analoghi uffici di Statistica furono istituiti in Piemonte nel 1836, in Toscana nel 1849, in Napoli nel 1851 ed in Roma nel 1852.

Riunite le varie parti della Nazione e proclamato il Regno d'Italia, con Regio Decreto del novembre 1861 fu fondata la *Direzione generale della statistica* alla dipendenza del Ministero di Agricoltura, Industria e Commercio; è in realtà abbastanza autonoma. Cesare Correnti ne fu il primo direttore; gli successe Pietro Maestri e dal 1872 sino al 1901 lo diresse Luigi Bodio; oggi la dirige il Comm. Aschieri. Presso la Direzione Generale della Statistica c'è un Consiglio superiore di Statistica — istituito nel gennaio 1887 — che raramente viene convocato, e una Giunta eletta in seno del Consiglio.

Il Consiglio superiore dà le norme generali per le ricerche da farsi; la Giunta elabora i metodi, vedendo quali si addicono di più alle ricerche speciali.

La Direzione generale è composta di due divisioni e di

parecchie sezioni. La prima divisione si occupa della popolazione e perciò del censimento, del movimento dello stato civile, delle condizioni sanitarie, ecc.—La seconda si occupa piuttosto della vita economica: della statistica industriale, dei prezzi, dei bilanci comunali e provinciali, ecc. — Molte delle attribuzioni della secondo sono passate all'*Ufficio di lavoro*. La statistica giudiziaria, che dipendeva dalla 1^a divisione nel 1908 è passata al Ministero di grazia e giustizia.

La Direzione della statistica fa molte pubblicazioni periodiche ed occasionali assai pregevoli.

La Direzione generale non ha veri suoi organi periferici; ma le funzioni raccoglitrice vengono esercitate, specialmente nei comuni, da funzionari locali alla dipendenza delle Amministrazioni municipali e con particolarità da quelli addetti allo stato Civile. In ciascuna provincia c'è una Giunta provinciale di statistica composta di otto membri e presieduta dal Prefetto; c'era una Giunta comunale di statistica, ma venne soppressa con decreto del 1887. Le sue attribuzioni passarono al sindaco. Le attribuzioni delle Giunte provinciali sono limitate ed è scarsa la loro attività; esse sostanzialmente sorvegliano le operazioni del censimento e talune inchieste speciali.

Oltre la Direzione generale di statistica, ciaschedun Ministero ha un ufficio suo proprio di statistica; sono bene organizzati quelli del Ministero del Tesoro, delle Finanze e che fanno periodiche ed interessanti pubblicazioni. Qualche volta i dati statistici speciali pubblicati da un Ministero sono diversi da quelli pubblicati dalla Direzione generale della Statistica. L'inconveniente è grave.

CAP. II.

OBBIETTO E DEFINIZIONE DELLA STATISTICA

7. *Definizione della Statistica.*—Si apprende dalla storia della Statistica come e quanto varia sia stata l'estensione del campo di azione della medesima: mentre gli uni (ad esempio Melchiorre Gioia) l'allargavano straordinariamente, altri la restringevano oltremisura e ne diminuivano il concetto e l'importanza.

Queste divergenze naturalmente si ripercuotono sull'*obbietto* da assegnare alla disciplina e sulla sua *definizione*.

E perciò che sulle definizioni della statistica c'è tutta

una letteratura essendosene enumerate oltre trecento, la maggior parte delle quali o mancano di precisione o sono monche, o sono unilaterali, o prolisse.

Fra le tante sembrami preferibile, sebbene io sia convinto della sua imperfezione, la seguente: *la Statistica è la disciplina dei fatti atipici, metodicamente raccolti in massa e ridotti in gruppi omogenei, della ricerca delle loro cause e delle loro leggi empiriche.*

Questa definizione: 1° elimina l'antica controversia sull'epoca cui deve riferirsi: se al solo presente o anche al passato: 2° nei fatti *atipici* comprende i fatti sociali — tra i quali stanno quelli politici — e gli altri appartenenti a fenomeni della natura: ad esempio quelli meteorologici; 3° tiene conto della sua caratteristica nel metodo che adopera: l'osservazione quantitativa: 4° indica la *riduzione in gruppi omogenei*, cioè in gruppi di fatti, che rappresentano tutti la stessa circostanza, cioè quel carattere di omogeneità necessaria per potere poi procedere allo studio delle cause e delle leggi dei fenomeni: 5° non accenna ai soli *fatti*, ma anche alla *ricerca delle cause* e alle *leggi*; e queste si denominano *empiriche*, per indicarne la natura, di cui si dirà in appresso: 6° infine si adatta alla statistica come *metodo* e come *scienza*; quantunque per me essa non sia che un *metodo*.

8. *Osservazione ed esperimento.* — In tutte le definizioni della statistica vi è l'accenno più o meno esplicito alle *osservazioni in massa e numerose*. Si deve spiegare questa circostanza notevole, ch'è propria della statistica, specialmente quando si occupa dei fenomeni sociali e demografici.

Ad intenderla meglio occorre premettere la distinzione tra *osservazione* ed *esperimento*, che alcuni confondono. Herschel giustamente osservò che l'*osservazione* e l'*esperimento* sono forme variabili dell'*esperienza* intesa in senso largo; e la prima chiamò *osservazione passiva* e la seconda *osservazione attiva*.

L'*esperimento* o *osservazione attiva* è la riproduzione artificiale delle condizioni, che producono, generano, un dato fenomeno. Il chimico mette in contatto gli atomi dell'ossigeno e dell'idrogeno e produce a sua volta l'acqua. Il fisico nel suo gabinetto artificialmente sviluppa l'elettricità. Il fisiologo ha a disposizione gli animali sui quali fa svariati esperimenti — taglia, ferisce, inietta, affama — onde studiare le funzioni dei vari organi, ecc. ecc.

L'*Osservazione passiva* può essere *interna* o *esterna*.

La interna, detta anche *introspezione*, è quella che l'uomo fa sopra sè stesso ed è del dominio della psicologia.

L'osservazione esterna—che può anche ricorrere all'uso di strumenti, che perfezionano, per così dire, i sensi (microscopii, termometri, ecc.) nelle scienze naturali e servirsi di loro accanto all'esperimento — quando si riferisce alla società umana, può essere *storica* o *statistica*. L'osservazione *storica* sia del presente, sia del passato, insegna ad accertare ed a conoscere, quanto più esattamente può, i fatti, descrivendoli come furono o come avvengono.

L'osservazione si dice *statistica*, quando le osservazioni sono fatte in massa, sono numerose ed esposte numericamente. Nella statistica il numero è non solo il mezzo, ma anche il limite delle osservazioni e dello studio dei fatti, poichè essa per lo appunto si occupa dei fatti suscettibili di espressione e notazione numerica. Coll'osservazione semplice o *passiva* non si producono artificialmente i fenomeni, ma si sorprendono e si studiano nelle loro naturali e spontanee manifestazioni. Non è pensabile, infatti, che il sociologo possa tentare di riprodurre artificialmente i fenomeni, che vuole studiare. Può riprodurre—come in *corpore vili*—le condizioni, che devono generare il furto, l'omicidio, il suicidio, ecc.? Può a sua volontà determinare i parti multipli o i sessi nelle nascite? No.

L'osservazione *passiva* deve limitarsi a notare i fenomeni demografici, sociali, meteorologici ecc. man mano che si presentano, e potrà meglio studiarli col *metodo statistico*, che sino ad un certo punto ne permette la misura. Le scienze che possono ricorrere allo esperimento o osservazione attiva, si dicono *sperimentali*. Invece diconsi *scienze di osservazione* quelle, che devonsi limitare alla semplice *osservazione passiva*. Sono brillanti e solide le prime; ma non si creda, che tutte le scienze di osservazione siano incerte ed in ugual modo. L'astronomia è una scienza di osservazione. Invece nelle scienze sperimentali non mancano incertezze e contraddizioni, che sono più numerose a misura che ci allontaniamo dalla materia inorganica: ad esempio, nella fisiologia.

La *statistica* e tutte le scienze sociali sono *scienze di osservazione* e non *sperimentali* per le ragioni esposte.

La *statistica* col suo metodo speciale e coll'elemento *numerico*, *quantitativo*, di cui si serve, all'operazione dà quella maggiore precisione, che non può venirle dalla semplice osservazione *storica*.

9. *Fatti tipici e atipici*. — La statistica deve fare le sue *numerazioni in massa*, perchè essa si occupa dei fenomeni o fatti *atipici*, che sono propri della medesima.

Un fenomeno, un fatto, dicesi *tipico*, quando basta l'osser-

vazione di un solo o di pochissimi casi o esemplari del medesimo per poterne stabilire i caratteri e intendere a quale genere o specie si appartiene quando si presenta alla nostra osservazione. In un gabinetto di mineralogia noi troviamo pochi esemplari o uno solo di ogni minerale, perchè, quando se ne conosce uno, si conoscono gli altri: la cristallizzazione, il colore ecc. del cloruro di sodio, dello zolfo ecc. sono identici in tutti i pezzi dello stesso minerale. Sono *tipici*.

Nella biologia questa *tipicità* diminuisce; pure, nelle grandi linee, la struttura di una foglia e del sistema vascolare osseo nervoso, ecc. di un animale di una data specie, di un dato genere, conserva abbastanza uniformità in tutti gli altri esemplari. E nell'uomo, guardato dal punto di vista esclusivamente anatomico o antropologico, c'è pure tale grado di uniformità.

Nel campo biologico diminuisce la *tipicità*, quando si studiano le funzioni degli organismi ammalati. Si attenua ancora di più nella meteorologia: un giorno dello stesso mese o della stessa stagione non si rassomiglia per le sue condizioni atmosferiche, per grado di calore, per la direzione ed intensità dei venti, ad un altro ecc. ecc.

La *tipicità* infine si riduce al minimum nelle manifestazioni della vita sociale. Mentre nel sistema osseo, venoso, ecc. di due uomini appartenenti alla stessa razza le variazioni saranno semplicemente di dimensione e presenteranno piccole differenze; le differenze saranno maggiori nelle manifestazioni morali, intellettuali ecc., connesse ai più complessi aggruppamenti di condizioni economiche, politiche, sociali. Il grado di intelligenza, di moralità, di attività, ecc. di un uomo sarà assai diverso da quello di un altro, specialmente se i due uomini, benché viventi nella stessa città e nello stesso tempo, appartengono a due classi differenti. Anche quelle manifestazioni, che più si avvicinano alla natura strettamente biologica, come, ad esempio, la natalità, la nuzialità — come espressione del bisogno sessuale, ecc. — sono diversissime da un gruppo sociale all'altro per le ragioni d'indole sociale, che si avrà occasione di studiare. I fatti, i fenomeni umani sono detti perciò *atipici*, perchè, quando avremo conosciuto un uomo nelle sue qualità intellettuali e fisiche, economiche e morali, non possiamo concludere che gli altri uomini siano identici o rassomiglino a quelli osservati.

In conseguenza, quando noi vogliamo studiare un fenomeno in una data collettività, dobbiamo studiare l'intera collettività. Dieci, venti, cento uomini osservati, potrà darsi che

non commettano alcun reato, non si suicidino, non emigrino, ecc.; per conoscere, quindi, il fenomeno *delitto, suicidio, emigrazione*. ecc., nelle sue proporzioni, noi dobbiamo studiare in massa tutti gli uomini, che fanno parte di quella data collettività. Perciò si dice che la statistica non si occupa dei fatti e fenomeni *singoli*, ma di quelli *collettivi*.

In tutti i fenomeni *atipici* rileviamo tutte le circostanze che li accompagnano, in guisa che poscia aggruppandole opportunamente possiamo cogliere il dato più *tipico* o più costante. Così, ad esempio, notando un primo, un secondo, un terzo, ecc. un ennesimo suicidio od omicidio, noi vedremo dalla rilevazione delle condizioni dei vari suicidi che il grado di coltura, la professione, la data età ed anche le stagioni in cui si verificano in cento, in mille casi prevalgono su tutte le altre condizioni. Queste circostanze, che prevalentemente accompagnano i suicidii, ne costituiscono il lato più *tipico*, più costante. Perciò si dice che mercè l'osservazione statistica dei fenomeni *atipici* si rileva quello che in essi vi è di *tipico*.

Quali che siano i fenomeni studiati dalla statistica, noi dobbiamo sempre tener presente che il suo metodo delle osservazioni multiple, se ci fa conoscere ciò che nei fatti e nei fenomeni, ha il carattere di costanza, e dà luogo alla formulazione delle *leggi*; queste *leggi*, però, non sono dei rapporti costanti, ma probabili, e non ci danno mai la ragione intima delle cose, che del resto non ci venne data sinora da veruna altra scienza. Essa, ad esempio, constata, che le morti maschili prevalgono sino ad una certa età sulle morti femminili; ma non sa dire perchè ciò avvenga. Inoltre essa ammaestra intorno al modo di usare i dati nello scioglimento di speciali quistioni, ma non deduce essa medesima dalle proprie leggi i principii direttivi applicabili ai casi particolari.

In modo generale, perciò, si può esprimere il compito della statistica in questi termini: scevera nei fenomeni collettivi ciò che vi ha di tipico nella varietà dei casi, di costante nella varietà, di più probabile nell'apparente accidentalità; decompone, fino al limite che la natura del metodo consente, il sistema di cause o forze, di cui essi fenomeni sono la risultante.

CAP. III

DEL METODO STATISTICO.

10. *La statistica è un metodo non una scienza* — Si è discusso lungamente se la statistica sia una scienza o un metodo.

Sono gravi le obiezioni contro la statistica considerata come scienza e si riducono principalmente a queste due:

a) Una scienza non si serve di un solo metodo. La statistica, come scienza investigatrice, non adibisce che il solo metodo *statistico*. Una scienza circoscritta all'uso esclusivo di un metodo riesce necessariamente monca, perchè le sfugge l'analisi di certi lati ed aspetti dei fatti studiati, per i quali sarebbe opportuna l'applicazione di altri metodi.

b) Ogni scienza ha un campo di fatti propri e di leggi proprie; la statistica invece raccoglie fatti, elementi, leggi in tutti i campi e li somministra a tutte le altre scienze politiche e sociali.

Questa è la obiezione più grave, che si può presentare contro la statistica come scienza. Poichè: o essa ha diritto di ricercare le cause e di formulare le leggi dei fenomeni, che studia, ed invade il campo di tutte le scienze sociali — in questo caso si confonde colla sociologia — ed anche di talune delle così dette scienze naturali: ad esempio, la meteorologia. O essa deve limitarsi a raccogliere e ordinare gli elementi per le altre scienze, ed in questo caso funziona esclusivamente come un metodo.

L'importanza grandissima del metodo non potrebbe bastare a farne una scienza a sè, autonoma, indipendente. La statistica, perciò, potrà essere considerata, oltre che come un *metodo*, anche come una *scienza* sino a quando del suo studio farà parte integrale la demografia. Ma la demografia ha già assunto tali proporzioni, che può prevedersene non lontano il distacco.

11. *Del metodo in generale*. — Tutte le critiche, che si rivolgono alla statistica come *scienza*, non hanno ragione di essere considerandola come *metodo*: e se anche le si neghi la dignità di *Scienza ausiliare*, assegnatale da Râmolin, è assolutamente innegabile la importanza di essa.

“ Per *metodo*, in generale, s'intende la via che segue il pensiero per scoprire la verità, e per comunicarla ad altri. „

Il *metodo*, perciò, è *inventivo* e *dimostrativo* o *didattico*. Il primo serve a scoprire la verità; astruendo dalle applicazioni esso riferiscesi ad ogni specie d'investigazioni; nelle applicazioni poi assume carattere specifico secondo l'oggetto di cui si occupa e gli strumenti di cui si serve. Il *metodo dimostrativo* serve a comunicare agli altri le verità scoperte.

Il *metodo*, secondo il suo punto di partenza, è *deduttivo* o *induttivo*.

Il *metodo deduttivo* si chiama anche *metodo razionale* o *a priori*, perchè discende per via di ragionamento da ipotesi e principii, comunque stabiliti, e va dal generale al particolare, dal composto al semplice. L'aritmetica, l'analisi algebrica, la geometria, le scienze filosofiche sono quelle che adoperano essenzialmente il *metodo deduttivo* e si dicono perciò *scienze deduttive*.

Il *metodo induttivo* si denomina anche *analitico* *metodo di osservazione*, *a posteriori*. Dai fatti risale ai principii, dal particolare al generale, dal semplice al composto, dai fenomeni particolari alle leggi generali.

L'induzione è l'operazione dello spirito per mezzo della quale noi inferiamo che ciò, che noi sappiamo essere vero in uno o molti casi particolari, sarà vero in tutti i casi, che rassomiglino ai primi sotto certi rapporti assegnabili. In altri termini, l'induzione è il processo pel quale noi concludiamo che ciò ch'è vero di certi individui di una classe è vero della classe intera, o ciò che è vero una volta lo sarà sempre in circostanze simili „ (*Stuart Mill*).

Così nella fisica si constata il fatto della tendenza dei corpi verso la terra e se ne induce la legge o forza di gravità; nella statistica si constata ad esempio, la proporzione dei nati morti legittimi e illegittimi e se ne induce la legge della maggiore mortalità intrauterina degli illegittimi.

La geologia, la metereologia ecc., sono scienze puramente induttive. Però in alcune scienze, stabilita una legge, un principio, per mezzo del *metodo induttivo*, per via del ragionamento, poi se ne deducono altre leggi e si fanno altre applicazioni. Nell'*ottica*, per esperimento, per osservazione di fatti si stabiliscono le leggi della riflessione, della trasmissione della luce; accertate tali leggi, col ragionamento matematico si deducono le soluzioni dei problemi della propagazione della luce attraverso i diversi mezzi. L'*economia pura* ha come base sperimentale il principio dell'egoismo — la così detta massima edonistica —, da cui deduce altre leggi ed altre applicazioni.

Il *metodo induttivo* si appoggia all'*osservazione semplice o passiva* e all'*osservazione attiva o esperimento* che si disse già in che cosa consiste. La *statistica* non si serve che della prima.

12. *Del metodo statistico.* — Il *metodo* si dice *statistico* o *quantitativo* o *numerativo* quando le osservazioni sono fatte in massa, numerose ed esposte numericamente, come fu rilevato precedentemente, o per motivi speciali derivanti dalle materie studiate.

Il *metodo statistico* nelle scienze sociali è il *metodo induttivo* per eccellenza, perchè ai caratteri generali del metodo induttivo aggiunge la quantità, la misura, per dare alle sue risultanze la maggiore precisione possibile.

I caratteri principali del *Metodo statistico* sono i seguenti:

a) È un metodo di osservazione semplice e non sperimentale, perchè esso considera e numera i fatti quali si presentano nelle loro manifestazioni, ma non ne muta le condizioni e i rapporti con altri fatti, come si richiederebbe per sperimentare.

Però, se muove dall'osservazione, rinunciando al vero e proprio esperimento, non si esclude che esso possa raccogliere i risultati di esperienze relative a fatti assai variabili, quali ad esempio, possono essere fatte dai biologi nelle loro ricerche sulla eredità dei caratteri mediante incroci artificialmente prodotti. Ma non è il metodo statistico che guida le esperienze; esso si limita a raccoglierne ed elaborarne i risultati.

b) Il metodo statistico è principalmente induttivo, perchè, osservando le manifestazioni di certi fatti collettivi, cerca in base ad esse di assurgere alle *leggi*, che si ritengono valere per la generalità del fatto e per i casi analoghi. Per es., la *legge* della mortalità in una popolazione si stabilisce sulle osservazioni di un numero rilevante di casi di morte; la *legge* poi si estende in via di approssimazione ai casi futuri e non osservati.

c) Il metodo statistico si differenzia dagli altri metodi di osservazione e di ragionamento, perchè la osservazione non si limita a poche manifestazioni di un fenomeno, ma al *maggior numero possibile* di esse, in modo da procedere alle osservazioni per *masse* o per *collettività*. In tal guisa riesce a trovare ciò che esiste di generale nei casi particolari e qual è l'ordine che presentano nel loro svolgimento e nella loro distribuzione. Non c'è un criterio assoluto per giudicare del numero dei casi, che si devono osservare: potrà essere pic-

colo per un fatto, che abbia poche manifestazioni; dovrà essere maggiore o minore, secondo la maggiore o minore *variabilità* individuale dei casi.

d) Il metodo statistico procede aggruppando ed ordinando sotto i più vari aspetti le singole manifestazioni del fatto collettivo. I fatti che cadono sotto l'osservazione statistica poichè complessi, sono ricchi di qualità e di caratteri vari oltre l'elemento fondamentale, che costituisce il fatto stesso. In una nascita, ad esempio, oltre l'elemento biologico di un organismo che sorge alla vita, vi sono altre qualità — il sesso, la legittimità o illegittimità, l'essere avvenuta in città o in campagna, da genitori di varia età ecc., ora tutti questi caratteri o aspetti, che differenziano un fenomeno collettivo nelle sue varietà e nei suoi modi di essere, vanno, ciascuno, osservati per masse, onde la necessità di ordinare e suddividere l'insieme delle osservazioni in altre masse, a seconda di quei vari aspetti, ossia di riunire i singoli casi osservati in *gruppi omogenei* tali, che i casi raccolti in ciascuno di essi abbiano certe qualità comuni — la legittimità o illegittimità, il sesso maschile e femminile, ecc.

e) Col metodo statistico compiendosi l'osservazione per masse, ossia enumerando, contando via via le singole manifestazioni individuali, i risultati a cui si giunge trovano la loro espressione concreta in un numero. Non ci si deve illudere sulla infallibilità o sul valore di questo numero: esso talora non corrisponde nemmeno a tutti i casi attraverso i quali si manifesta un determinato fenomeno e che sarebbe stato conveniente osservare. Per esempio, non tutti i reati che si commettono, ma solo quelli *denunziati* all'autorità giudiziaria sono compresi nelle statistiche criminali e parecchi — specialmente truffe, frodi, reati contro il buon costume — sfuggono alla nostra conoscenza. Però coll'elemento numerico possiamo misurare in un modo più preciso, che altrimenti non potrebbe farsi, la entità e le variazioni del fenomeno e ad esso si permette l'uso di metodi matematici così di correzione come di analisi e di sintesi.

f) Il metodo statistico permette che si eserciti sulle masse osservate il ragionamento logico per via di *comparazioni*, al fine di determinare il modo di essere e la legge del fenomeno. La *comparazione*, rilevando le variazioni dei fatti da tempo a tempo, da luogo a luogo, in queste od in quelle circostanze, agevola a trovare le cause e le leggi. Così, osservando che la mortalità da qualche tempo è in diminuzione in vari paesi ed osservando che tale diminuzione è ac-

compagnata da certe circostanze economiche, demografiche, igieniche, ecc., tali circostanze si assumono come *cause* della diminuzione stessa (*Bosco*).

Su questi caratteri del metodo statistico ritorneremo discutendo delle singole parti della metodologia.

CAP. IV.

IMPORTANZA E UTILITÀ DELLA STATISTICA.

13. *Importanza scientifica e politico-sociale della Statistica.*— Siano quali che si vogliano le opinioni degli scienziati sui punti più controversi sinora discussi, non si può negare che la statistica abbia acquistato una importanza considerevole.

Dal punto di vista scientifico si ricordi che la statistica va dando sempre più base positiva alla economia politica, alla sociologia criminale, alla meteorologia ed anche alla biologia. Nè sono minori i servizi che rende allo Stato e alla società sotto l'aspetto pratico.

Come potrebbe per esempio, il Governo fissare o ripartire giustamente il contingente militare, senza la statistica della popolazione? Come fissare o ripartire, con equità, le imposte, concludere un trattato di commercio, stabilire o riformare una tariffa interna, se non si volge, come a punto di appoggio, oltrechè alle inchieste propriamente dette, anche alla statistica economica?

E basta guardare alla colossale importanza, che hanno assunto gli *Uffici di statistica del lavoro* in America e in Europa, pei servizi che essi rendono alla legislazione sociale, per comprenderne la importanza acquistata e la immensa utilità.

Ed è la statistica che ha dato agli economisti ed ai sociologi delle consolanti sorprese: i suoi dati hanno insegnato che possono diminuire le ore di lavoro ed aumentare la produzione del lavoratore, che può elevarsi il salario e diminuire il costo del lavoro, ecc. Ciò risulta dalle ricerche di Brassey, di Gavernitz e soprattutto da quella grande inchiesta di Gould, che Carrol-Wright chiamò *il trionfo della statistica*.

Al Governo ed ai privati è la statistica dei trasporti che insegna se tale ferrovia, tale canale, devono essere costruiti e quali somme è ragionevole consacrarvi. È la statistica della popolazione che insegna quale proporzione dobbiamo dare ad un ospedale, ad una scuola ecc. È la *statistica* che, special-

mente nelle nuove, insegna alle industrie su quali vendite possono contare. E la *statistica* che in tutto il campo della produzione regola le importazioni o le esportazioni, e si studia di evitare le crisi acute e gravi e che dai *trusts* d'interesse privato forse condurrà ai *trusts* nell'interesse collettivo. Le compagnie di assicurazione non vivono che di statistica. La statistica, specialmente agricola, è indispensabile al commercio cui insegna se il raccolto è deficiente in un tale paese e se in un altro è abbondante; e dalla mancanza di buone statistiche deriva la instabilità del commercio.

Sulle ricerche e sulla pubblicazione rapida di statistiche esatte è fondato l'*Istituto internazionale di agricoltura*, che si deve alla iniziativa di Lubin o del Re d'Italia.

Riassumendo: la statistica è necessaria perchè, per sviluppare la ricchezza di un paese, bisogna conoscerne la sorgente; per soddisfare i bisogni, si deve conoscerne la estensione; per estendere il commercio, bisogna conoscere gli sbocchi; per aumentare o frenare la popolazione, bisogna conoscere le leggi che la regolano; per decidere l'esecuzione di un'opera di utilità pubblica, bisogna sapere in quale misura sarà profittevole, ecc.

La statistica è necessaria all'industriale, al commerciante all'operaio, allo scienziato, al filosofo. La statistica soprattutto è il mezzo di controllo del funzionamento delle pubbliche amministrazioni e dell'azione dello Stato.

Comunque, la statistica si vedrà meglio apprezzata quando seguirà più sistematicamente un indirizzo pratico, che possa renderla sempre più utile agli statisti ed ai politici.

Parimenti, quanto più accurate saranno le ricerche statistiche, quando più diligente sarà la elaborazione e la esposizione dei dati raccolti, quanto più prudenti e modeste le sue conclusioni, tanto meno potranno trovare pretesti e ragioni di muovere critiche ingiuste o ragionevoli i nemici e gli scettici della statistica.

14. *Differenze tra la statistica e altre discipline.* — Considerando la statistica come un *metodo* e non come una *scienza* non c'è più bisogno di segnalarne i rapporti e le differenze colle altre scienze e di stabilirne i limiti precisi. Basterà, quindi, dir brevemente dei caratteri, che la differenziano dalla matematica, dalla storia e dalla sociologia, con le quali taluni intimamente la connettono.

a) Il metodo statistico, o la statistica puramente e semplicemente, si distingue dalla matematica non solo pel suo carattere strettamente metodologico, ma per la natura del

suo oggetto, consistente non già in quantità astratte, ma in fatti tradotti in quantità concrete. Perciò chi si avvisasse di chiamare matematico il metodista-statistico, perchè egli pure fa uso di segni simbolici (cifre numeriche e figure geometriche), cadrebbe nel medesimo errore di chi reputasse matematico il cassiere per ciò solo che numera.

b) La distinzione più difficile a stabilire è quella colla storia.

Storia e statistica differiscono essenzialmente pel metodo: in quanto la prima segue lo svolgimento dei fatti, ne indaga le cause e ne descrive gli effetti; mentre la statistica rappresenta ordinati i gruppi omogenei quantitativi.

La statistica è il *nosce te ipsum* applicato alle nazioni; ma la statistica non si può adoperare sempre e allora si ricorre ad un'altra forma tipica d'induzione, meno precisa ma non meno importante: la osservazione storica. Perchè la *storia* ci spiega i fatti della società nel loro ordine di successione, mentre la *statistica* ne dimostra le relazioni di coesistenza; l'una serve a chiarire il nesso causale o la dipendenza reciproca dei fenomeni come si svolgono nel tempo; l'altra come si attuano nello spazio; ed entrambe ci forniscono la materia concreta degli studi sociali e la base delle investigazioni scientifiche (*Chevalier*). La storia ci mostra il concatenamento degli avvenimenti; ma essa sola è insufficiente, perchè, se essa mette in mostra gli attori apparenti, dimentica altre specie di fattori, che sono più importanti: un cattivo raccolto, una crisi commerciale. ecc. esercitano una influenza maggiore della volontà o dei capricci di un Re, di un Imperatore, di un Presidente. La statistica è meno compiuta della storia in quanto non dà l'identità delle persone, ma le considera come soggetti di numerazione; non fa conoscere le peculiarità caratteristiche delle singole unità e non ci dimostra se queste sono buone o cattive; mentre la storia ci dà il dramma vero. Nella storia campeggiano gli *eroi*; nella statistica le *collettività* anonime.

Sarà bene avvertire, però, che oggi fra gli storici si è determinata una corrente, che tende ad attenuare le differenze, già incerte, fra storia e statistica, volendosi che la prima si valga meno del lato biografico e più delle manifestazioni collettive.

A questa nuova corrente hanno impresso un vigoroso impulso coloro che fanno capo al *determinismo economico*; ma l'esagerazione della scuola è evidente e contro di essa comincia la reazione. Si torna a riconoscere che negli avveni-

menti umani esercitano un'azione gli *uomini rappresentativi* di Emerson, gli *eroi* di Carlyle — Socrate, Cristo o Mazzini; Cesare, Carlo Magno o Napoleone I.; Washington o Garibaldi — e con questa azione ammessa nella narrazione concatenata degli avvenimenti e colla grande prevalenza del metodo descrittivo e dell'apprezzamento *qualitativo* nell'una, del metodo numerico e dell'apprezzamento *quantitativo* nell'altra, torna a stabilirsi una diversificazione sufficiente tra la storia e la statistica.

Rimarrà la differenza fra le due discipline; ma è evidente che la statistica diverrà sempre più necessaria allo storico.

c) Intimissimi sono i rapporti tra la Statistica e la Sociologia *. Li constatai sin dal 1891, inaugurando il mio Corso libero di Statistica nella Università di Palermo, ed appresi dopo, che il Neumann-Spallart aveva fatto oggetto di una delle sue pubblicazioni questo rapporto intimo tra Statistica e Sociologia.

Si deve al von Mayr una specie di confusione tra *statistica* e *sociologia*, avendo chiamato la prima pretensiosamente una *sociologia esatta*. La confusione non è più possibile lasciando alla *Statistica* il solo carattere metodologico; ma anche lasciandole la dignità di scienza e comprendendo nella sua trattazione la demografia, la *Statistica* non sarebbe sempre che una parte della *Sociologia*.

La Sociologia, infatti, sintesi o risultante di tutte le scienze sociali, si serve della psicologia della etnologia, della storia, del diritto, della economia, della politica, della demografia, ecc. non può quindi confondersi la parte col tutto e non si deve accettare, per evitare equivoci, che possono anche influire sull'indirizzo delle ricerche, la denominazione di *Sociologia esatta* assegnata dal von Mayr, come si disse, ed anche dal Salvioni.

* Le definizioni sono sempre difficili; difficilissima quella della *Sociologia*. Alcuni semplicisticamente la definiscono: *la scienza dei fenomeni sociali*. Meglio la si intende definendola: *la scienza della struttura e della vita delle società umane*.

LIBRO II.

STATISTICA METODOLOGICA

A) Rilevazione.

CAP. I.

RILEVAZIONE DIRETTA.

15. *Partizioni del metodo statistico.* — Divido il *metodo statistico* in quattro parti, come fa il Ferraris per la statistica teorica generale e tecnica, e cioè: 1.º *Rilevazione*; 2.º *Elaborazione*; 3.º *Esposizione*; 4.º *Interpretazione*. Gli avvedimenti speciali, che si devono adoperare nelle particolari applicazioni della tecnica statistica, devono essere trattati nelle singole parti della statistica applicata (demografia, statistica morale, statistica economica, ecc.).

Si avverte, che nel considerare *quantitativamente* un qualunque fatto collettivo la nostra osservazione può dirigersi in vari modi: a) Possiamo misurare il fatto nella sua entità numerica; per esempio, quanti i gradi della temperatura, quale il numero dei morti in una data popolazione, quale il saggio dei salari. — b) Possiamo misurare il fatto nelle sue variazioni; per esempio, le mutazioni della temperatura nelle varie ore del giorno, nei giorni o nei mesi dell'anno, le variazioni mensili o annue delle morti, dei salari, ecc. — c) Possiamo misurare il fatto nella sua distribuzione, ossia misurare il modo con cui le molteplici manifestazioni individuali si raggruppano in uno o in un altro senso, talora senza alcuna norma apparente; tal'altra con un certo ordine e regolarità; per es., come i gradi della temperatura nel corso di un giorno, di un anno, ecc., salgono e si portano da certi massimi a certi minimi, come i morti si dispongono secondo le varie età, come i salari si addensano intorno ad un certo saggio, da cui re-

stano più o meno lontani i salari singolarmente scarsi o remunerativi „ (Bosco).

16. *La rilevazione.*—Dicesi *rilevazione* la raccolta ad una ad una delle manifestazioni del fenomeno con tutte le circostanze, che lo accompagnano in un determinato momento del tempo e dello spazio.

La riunione delle manifestazioni del fenomeno costituisce il *dato statistico*. Un *dato* individuale non è il vero *dato statistico*; il *dato* viene costituito dalla massa delle manifestazioni di un fenomeno che si sono manifestati in un determinato tempo e in un determinato spazio. In questo senso, per es., si parla dei *dati* sulle nascite, sulle morti, sui matrimoni ecc., mentre ogni nascita giornaliera ecc. è un *elemento*, una *unità*, che serve per costituire il *dato* sulle nascite annuali o mensili, ecc.

Non si rilevano i *dati*, ma gli *elementi* o *unità* che lo costituiscono.

La *rilevazione* dicesi *diretta immediata*, quando le manifestazioni del fenomeno si rilevano direttamente nella loro qualità e nella loro totalità.

La *rilevazione* dicesi *indiretta* quando l'osservazione non cade direttamente sulle manifestazioni del fenomeno, che si vogliono rilevare o quando le manifestazioni non si rilevano nella loro totalità. Questa rilevazione costituisce la cosiddetta *statistica congetturale*.

Il *dato statistico* può essere positivo o congetturale; originario o derivato.

E' *positivo*, quando esprime un numero certo e definito ottenuto mercè la rilevazione diretta e nella sua integrità; è *congetturale*, quando è soltanto approssimativo ed ottenuto per mezzo della rilevazione indiretta.

Si dice *originario*, quando viene presentato tale e quale è stato ottenuto dalla rilevazione: ad esempio: il numero assoluto dei nati, dei morti in un dato anno; la popolazione assoluta di un dato paese rilevata col censimento.

Dicesi *derivato*, se il *dato* è stato reso proporzionale ad un altro fenomeno o ad un dato numero; ad esempio: il numero dei nati, dei morti di un dato anno, è stato proporzionato a 1000, a 100mila abitanti; il numero degli abitanti ad una data superficie — ordinariamente, ad un chilometro quadrato — di un dato paese.

Il *dato* infine sarà *numerico*, se espresso da *numeri* o *cifre* e in forma rigorosamente definita per quanto concerne la quantità. Solo in via subordinata, talora al difetto del nu-

mero si supplisce con altre espressioni quantitative. E' *descrittivo*, se espresso colle parole. Si discute molto se si dovevano anche usare le parole. Ora si usano quasi sempre contemporaneamente le une e le altre. La *qualità* non potrà indicarsi che colle parole.

Nella *rilevazione diretta* si deve distinguere: a) *quali dati* si devono rilevare; b) il *tempo* e il *luogo* delle rilevazioni; c) il *modo* e la *specie* delle rilevazioni; d) *chi* deve rilevare.

17. a) *Analisi qualitativa*. — Anzitutto si deve porre mente a ciò che si deve rilevare; perciò l'*analisi qualitativa* deve precedere lo studio *quantitativo* dei fenomeni.

Nell'*analisi qualitativa*, che serve per conoscere ciò che si deve rilevare, c'è bisogno del sussidio delle altre scienze. Così se si devono rilevare i *dati* sui *salari*, sui *prezzi*, ecc., è l'economia che darà l'idea del salario, del prezzo, ecc.

L'*analisi qualitativa* serve per classificare i dati da rilevare; quella *quantitativa* per misurare l'intensità della manifestazione del fenomeno.

Si avverta che bisogna *definire* uniformemente il *dato* per renderlo comparabile; si deve *definirlo* esattamente, affinché ciascun fatto rimanga distinto da ogni altro fatto, anche affine, e che i risultati della rilevazione, riferendosi tutti allo stesso oggetto, riescano *omogenei* e possano essere riuniti in gruppi. La *definizione* esatta non è sempre facile. Spesso ci dobbiamo servire della definizione giuridica, così, per sapere chi sia l'*emigrante*, si risconterà la legge del 31 gennaio 1901; quale sia la *popolazione legale* e la *popolazione di fatto* si desumerà dalle leggi sui Censimenti.

Ma spesso la definizione è diversa, secondo l'ufficio che fa la ricerca e lo scopo per cui si fa. Così a Parigi, secondo il censimento, nel quartiere Berry, vi sono 305 *case*, perchè tanti sono gli ambienti permanentemente abitati; per l'ufficio delle imposte, invece, vi sono 1214 *case*, quanti sono gli ambienti sottoposti ad imposta, destinati non solo ad abitazione delle famiglie, ma anche a botteghe, a *bureaux*, ecc., (*Bertillon*).

La *quantità* è bene si sappia che talora nella *statistica* diviene *qualità*. Come dal numero delle vibrazioni nell'*acustica* si forma la *qualità* del suono: alto, medio, basso; così nella *statistica* dal numero dei morti si deduce una *qualità*: la *mortalità* di una data professione, di un dato ambiente, ecc.

In quanto alla *qualità* delle manifestazioni del fenomeno da rilevare il *dato* deve essere:

1.° Suscettibile di essere ridotto ad *espressione numerica*. L'estetica, la religione non sono suscettibili di essere ridotte ad espressione numerica.

2.° *Utile o importante* dal punto di vista scientifico, amministrativo o sociale. Non c'è alcun dubbio sulla utilità di una statistica dei salari, dei prezzi, dei trasporti, ecc. I criteri della utilità e della importanza sono mutabili e relativi.

3.° *Opportuno*: cioè la spesa ed il tempo per la rilevazione devono corrispondere all'utilità dei risultati. Così la statistica della spedizione delle lettere, ecc. se si facesse giorno per giorno, importerebbe una spesa ed un lavoro sproporzionati ai risultati; perciò si fa soltanto in una settimana dell'anno—e si sceglie quella, che corrisponde al movimento medio dell'anno.

4.° *Conveniente*: non si devono offendere le convinzioni— forse si potrebbe dire: i pregiudizi — dei cittadini. Per ragioni di convenienza, nel censimento italiano del 1881, si omisero i *dati* sulla religione. Non è *conveniente* richiedere un *dato* su di una quistione, alla quale rispondendo in un modo o in un altro, si può assicurare o perdere un beneficio.

Ad esempio, nel censimento degli Stati Uniti del 1890 si domandò se si era stati soldati o marinai in tempo di guerra. La domanda sembrava fatta allo scopo di dare o no una pensione (*Mayo-Smith*).

5.° *Veridico*: una qualità, che non ha bisogno di essere illustrata. Per averlo veridico quanto più è possibile si deve guardare alla origine del dato, alla condizione di chi lo raccolse, ecc.

6.° *Preciso*: per certe manifestazioni la precisione si può dire ch'è massima: *nascite, morti, matrimoni*; è molto minore per parecchi altri: *migrazioni, reati*, ecc.

7.° *Adeguato*: cioè non deve comprendere nè di più nè di meno di quello, che si vorrebbe che esso comprendesse; deve corrispondere alla manifestazione reale. Si badi che allora il dato può essere incompleto, senza essere erroneo.

Nelle statistiche penali, il *dato* delle condanne è esatto; ma è inadeguato, se lo si vuol prendere qual espressione generale della criminalità.

È grandissima la relatività di tutte le condizioni del *dato statistico*. Tutto ciò, che oggi sembra inutile, un tempo parve cosa utilissima.

Rilevando il fenomeno, si disse che dovevasi prendere nota

di tutte le circostanze connesse col medesimo, per vedere se le sue manifestazioni hanno identità apparenti o sostanziali. Per esempio: nello studio dell'*emigrazione* occorre distinguere se è *temporanea* o *permanente*; se vi danno maggiore contributo i contadini o gli operai, i celibi o i coniugati, i giovani o gli adulti, gli uomini o le donne. Tutte queste circostanze danno al fenomeno una importanza ed un carattere diverso. Per le nascite si deve osservare se sono legittime o illegittime, ecc. Perciò, dopo che si è assegnato il *genere* del fenomeno, bisogna per quanto è possibile determinare la *specie*. Nel genere *emigrazione* si deve distinguere la specie: *temporanea* o *permanente*; nel genere *nascita*, la specie: *legittima* o *illegittima*, ecc.

S'intenderà per *genere* quella unità collettiva di fenomeni, che contiene altre unità collettive sotto di sé; le unità collettive contenute nel genere si diranno *specie*; incontreremo perciò nella statistica fenomeni generali e fenomeni speciali. Diremo fenomeni generali quelli, che contengono altre categorie di fatti sotto di sé (per es., nascite); diremo fenomeni speciali quelli, che sono contenuti nei fenomeni generali (per es. nascite *legittime* o *illegittime*).

Del resto è difficile distinguere i fatti assolutamente generali dei fatti assolutamente speciali, perchè questi ultimi contenuti in un fatto generale, possono alla loro volta contenere altri fatti sotto di sé, cosichè possiamo avere più gruppi discendenti da un fatto generale.

I fatti *speciali* statistici, che entrano a costituire un fatto generale, possono essere:

a) d'indole diversa fra loro, ma subordinati ad uno stesso fatto generale, per es. le nascite, i matrimoni, le morti, compresi tutti nel fenomeno generale: *movimento della popolazione*.

b) d'indole affine tra loro, ma distinti soltanto per qualche elemento caratteristico, per es. nati vivi e nati morti, morti maschi e morti femmine;

c) d'indole identica tra loro, ma distribuiti in tempi e luoghi diversi — così i dati sui salari, da tempo a tempo, da luogo a luogo, formano un genere distinto appunto per vicende di tempo e situazione di spazio. Infatti, per potere procedere all'enumerazione o analisi quantitativa, occorre che vi sia una separazione, qualunque essa sia, fra i fatti omogenei, che vi si assoggettano. Il tempo diverso e lo spazio diverso rappresentano appunto una delle forme di *diversità* e quindi di possibile rilevazione numerica (*Salvioni*).

18. b) *La rilevazione in rapporto allo spazio.* — A meglio qualificare i fenomeni, si deve avere riguardo allo *spazio* ed al *tempo*, nel quale essi devono essere rilevati. Il tempo e lo spazio sono funzioni dei fenomeni nel senso che questi si manifestano in un certo *spazio* e in un certo *tempo*; i fenomeni alla loro volta, sono funzioni dello *spazio* e del *tempo* perchè destinati a mutare col mutar di quelli.

Cominciando dallo *spazio* si deve porre il quesito: quale estensione devono prendere le indagini statistiche? In astratto si può rispondere, che, come per gli oggetti, così per lo spazio, la statistica mira all'universalità; perciò tali indagini dovrebbero e potrebbero abbracciare tutto il mondo.

Ma a questo sconfinato ampliamento delle ricerche si oppongono limiti di *fatto* e limiti d'indole *teorica e scientifica*.

Di *fatto*. 1.º Tutto il mondo non può diventare campo di osservazioni statistiche.

2.º Vi sono fatti d'indole politica che prendono la loro forma dal diritto pubblico dello Stato e che quindi non consentono l'estensione delle osservazioni fuori del medesimo; ad esempio: i reati.

3.º Vi sono fatti che non sono *dispersi* o non si verificano per tutto il territorio, ma di loro natura sono *concentrati* in determinati spazii, ad esempio: gli studenti universitari.

Le rilevazioni si fanno per provincie, circondari, comuni, in Italia; per dipartimenti, circondari, comuni in Francia; e analogamente, altrove.

Questo frazionamento amministrativo però non è valevole, specialmente dal punto di vista scientifico delle indagini statistiche. Invero il confronto dei risultati fra luogo e luogo diventa difficile, sia che si guardi alla diversa *estensione* territoriale delle singole circoscrizioni amministrative, sia alle diverse *condizioni* principalmente *fisiche* dei territorii, che possono essere paludosi, boschivi, piani, montuosi umidi, ecc. Epperò l'adottare per le ricerche statistiche le divisioni dello spazio di carattere politico e amministrativo è inevitabile ma se il fondamento della statistica è la *comparazione*, i risultati, che si ottengono in siffatta guisa, sono poco comparabili — assai meno di quelli relativi alle divisioni del tempo.

Le osservazioni, che si fanno per parecchi anni in uno stesso luogo, si possono confrontare agevolmente, perchè gli anni (e coll'anno i mesi, i giorni, ecc.), rappresentano intervalli identici di tempo universalmente conosciuti.

Non sono confrontabili invece i risultati greggi di una ri-

cerca statistica operata in circoscrizioni diverse, ad esempio; la provincia di Cagliari: che ha 13,483 chilometri quadrati non è comparabile con quella di Napoli che ha chilometri quadrati 908.

In quanto all'inconveniente della diversa estensione, si ripara proporzionando il fenomeno alla superficie: così, per la popolazione, se ne determina la *densità*, cioè si proporziona il numero degli abitanti delle due provincie colla superficie, e si viene a conoscere che Cagliari secondo il censimento del 1911, aveva 38,7 abitanti per kilometro quadrato, mentre ne aveva 1443,6 la provincia di Napoli.

Con tale espediente, però, non si eliminano gli inconvenienti derivanti dalla difformità delle condizioni qualitative dei territori delle due provincie. A quest'uopo il Mayr consigliò il *metodo*, che egli disse *geografico*, che consiste nel raccogliere i dati statistici sopra menome porzioni del territorio, nello sminuzzarlo, per ricavare le condizioni di fatto delle molecole che lo compongono e poi coll'assimilare fra loro quelle che rappresentano risultati affini, ricomporre il territorio giusta vere *zone geografiche naturali* per ogni singolo fenomeno umano ed ottenere le zone di geografia antropologica, come si hanno quelle della geografia vegetale ed animale. Un modello vero di applicazione del *metodo geografico* si ha nell' *Antropometria militare* del D.r Rodolfo Livi, in cui tutti i dati antropometrici dei soldati italiani — statura, colore della pelle, degli occhi, dei peli, ecc. — sono raccolti, non per provincia o per circondario, ma per mandamento.

È stata eccellente l'applicazione, perchè i mandamenti in Italia appartengono di ordinario ad un territorio analogo per configurazione e situazione geografica, per la qualità del suolo pel clima, per la stirpe che li popola.

Il *metodo geografico* si presta meglio di ogni altro alla *ricerca causale*, pur non potendosi rassomigliare, come vorrebbe il Mayr ad una specie di esperimento statistico.

È difficile determinare in modo generale l'unità di superficie da adottare pel *metodo geografico*; si dice in generale: *devono essere le più piccole possibili*. Ma non troppo piccole, perchè altrimenti farebbe difetto *la massa delle osservazioni*. Il Mayr vorrebbe, perciò, assegnare a queste circoscrizioni una superficie di circa 500 km. quadrati. Si comprende agevolmente, però, che questa superficie è molto relativa, dipendente soprattutto dalla varia *densità* della popolazione.

La rilevazione in ordine allo spazio in conformità delle

circoscrizioni politiche, amministrative, giudiziarie, ecc., in Italia può essere *nazionale*, *compartimentale*, o per *regioni*, per *provincia*, per *comune*, per *distretto di Corte d'Appello*, ecc. Analogamente, presso gli altri Stati.

19. *La rilevazione in rapporto al tempo.* — I dati si manifestano in un dato tempo e in un dato luogo; non si manifestano una volta sola, ma si ripetono, si riproducono nel tempo. Perciò la osservazione statistica di un fatto non deve restringersi ad un solo momento, ma estendersi indefinitamente nel tempo, perchè altrimenti non potrebbe essere soddisfatta la esigenza della scienza statistica, che mira a scoprire la *legge del movimento del fatto*.

Sarà diversa la rilevazione, secondo che si tratti di fenomeni, che basta rappresentare in un determinato momento del tempo, in una condizione di *stato*; ovvero di fenomeni, che si devono seguire, almeno per un certo tempo, nel loro svolgersi successivo, osservarli cioè in una condizione di movimento. La distinzione è relativa, non assoluta: lo *stato* di un fenomeno non è, che un movimento colto in un punto della sua curva, e il *movimento* non è altro, che una successione degli stati del fenomeno.

Pei fenomeni, di cui si rileva la situazione in un dato momento, importa la scelta di questo. Il momento dev' essere tale da potersi accertare, nel miglior modo, i caratteri propri del fenomeno.

La scelta del momento più opportuno dipenderà quindi dalla natura del fenomeno e dagli intenti della rilevazione.

I momenti, nei quali la *rilevazione* vuol essere rinnovata, dipendono dalla natura dei fatti, che ne sono l'oggetto.

Vi ha fatti, che si succedono l'uno all'altro quasi senza interruzione od a brevi intervalli. Ve ne sono altri, invece, i quali si succedono modificandosi ad intervalli più o meno lunghi. Perciò in ordine al tempo la rilevazione può essere *continua*, *periodica* ed *occasionale*.

La rilevazione *continua* è quella che si riferisce ai fenomeni, che produconsi in ogni *unità* di tempo; morti, nascite, reati, ecc. E necessario di adottare una uniforme *unità di tempo*, perchè la rilevazione abbia carattere di continuità sistematica e possa seguire l'osservazione del fenomeno ad uguale distanza.

L'*unità di tempo*, secondo il fenomeno che si studia, può essere l'*ora*, il *giorno*, la *settimana*, la *decade*, il *mese*, l'*anno* ecc.

L'intervallo, che deve correre tra una rilevazione e l'altra deve stabilirsi in rapporto alla variabilità maggiore o minore

del fatto ed all'interesse, che può avere la conoscenza di queste variazioni.

Perchè, ad esempio, sia utile l'osservazione della temperatura nella febbre, le constatazioni devono essere fatte in diverse e determinate ore del giorno, perchè nella tifoidea si sa che in certe ore la temperatura si eleva e in altre si abbassa. L'osservazione del decorso della temperatura permette in questo caso di fare la diagnosi esatta.

Si dice *periodica* quando si ripete ad intervalli; come ad esempio il *censimento*; si denomina *occasionale* quando si fa una rilevazione in determinate circostanze, come ad esempio quella dei beni appartenenti alle opere pie, ecc.

20. c) *Modi e modelli della rilevazione.* — La rilevazione può essere *automatica* o *riflessa*. L'*automatica* si dice anche *spontanea*: è la più perfetta; è continua; vi sono persone espressamente addette, che registrano i fenomeni man mano che si presentano; gl'individui, per lo più, sono obbligati a denunziare per legge il fenomeno di cui essa si occupa. I registri per lo *Stato civile*, i registri *doganali*, *giudiziarî* ecc., costituiscono gli esempi più noti e chiari della rilevazione *automatica*.

Colla rilevazione *automatica* si ha quasi un trasporto nel campo sociale di quel sistema meccanico, che si usa negli osservatori metereologici, dove tutti i dati relativi alla temperatura, all'altezza barometrica, vengono segnati meccanicamente da strumenti speciali, senza che se ne occupino direttamente gli addetti all'ufficio.

La rilevazione *riflessa* avviene a periodi irregolari, occasionalmente; può essere anche periodica. Il censimento è una rilevazione *riflessa*. Si va in cerca del fenomeno, che si vuole studiare; e lo si studia di ordinario per mezzo delle inchieste.

Gli *strumenti materiali* per la rilevazione o *modelli di rilevazione* più comuni sono i seguenti: *registro*, *lista*, *bollettino individuale*, *tavola*, *questionario*.

Nel *registro*, ch'è un complesso di fogli o fascicolo o libro, notansi i fenomeni man mano che si verificano: *registro dello stato civile*, ecc.

Nella *lista* si aggruppano più unità statistiche, che costituiscono un ente collettivo. Ad esempio, la *lista di famiglia* nel censimento italiano del 1881, che in apposite *colonne* registrava le notizie sul *sex*, *età*, *stato civile*, ecc.; dei membri della famiglia.

Nel *bollettino individuale* si registrano le notizie concer-

nenti una sola unità statistica. Così, si raccolgono i dati sulla età, sesso, stato civile, istruzione, professione, ecc., di ogni singolo delinquente.

La *tavola* contiene diverse colonne, intestate secondo le diverse condizioni del fenomeno o che sono assegnate a diversi fenomeni.

Nella *tavola* per la emigrazione, in una colonna si nota il sesso dell'emigrante, in un'altra l'età, la professione, ecc. Del questionario si dirà tra poco.

21. *d) Chi deve rilevare i dati.* — Eccoci ai cosiddetti *organi della rilevazione*. Queste persone, o organi della rilevazione, sono *pubbliche* o *private*.

I singoli *privati* talora raccolgono notizie statistiche su di un dato fenomeno, su di una data collettività. In Italia sono notissimi i libri e le inchieste di Sonnino e Franchetti sulle *condizioni de' contadini e sulle condizioni politiche e sociali* della Sicilia e del Mezzogiorno. E' celebre la grande inchiesta di Charles Booth *sulla vita e sul lavoro del popolo* in Londra. I privati talora si riuniscono in associazioni. Tutte le rilevazioni dei privati sono saltuarie, occasionali, incerte.

La rilevazione di ordinario è *pubblica* ed *ufficiale* e si eseguisce per mezzo d'impiegati permanenti o temporanei alla dipendenza dello Stato, dei Comuni, delle Provincie, ecc. E' preferibile per la regolarità ed esattezza dei dati, che in gran parte deriva dalla pratica e dalla imparzialità di coloro che li raccolgono. Lo Stato ha il diritto e il dovere di far fare le rilevazioni necessarie pel suo funzionamento ed utili al corpo sociale. Gli *organi pubblici* si distinguono in: *dirigenti*, che stanno al centro; e raccoglitori, che stanno all' periferia.

Delle persone, che raccolgono i dati, dice il Messedaglia che esse devono: *sapere, volere e potere*. Cioè: devono conoscere bene la materia, ch'è l'oggetto della rilevazione statistica; pur sapendo eseguire la rilevazione, non devono essere impediti da ostacoli morali o intellettuali; sapendo e potendo, devono sentire il dovere di eseguire onestamente la rilevazione.

CAP. II.

RILEVAZIONE INDIRETTA O STATISTICA CONGETTURALE

22. *Perchè si ricorre alla rilevazione indiretta.* — Quando non è possibile la rilevazione immediata, diretta e precisa, di

cui ci siamo occupati, si ricorre alla cosiddetta rilevazione indiretta, mediata, che forma la *statistica congetturale*.

Si ricorre a questo metodo;

a) quando la spesa e la fatica per fare l'indagine diretta, esatta, non corrispondono all'importanza del risultato;

b) quando i fenomeni da rilevare, per la loro sede e vasta estensione, presentano quasi l'impossibilità di essere sottoposti a rilevazione diretta. È il caso della statistica della produzione agraria;

c) quando i fenomeni per la loro natura non sono suscettibili di una misurazione diretta. Si può ad esempio, tradurre direttamente in cifre la moralità di un uomo o di una società?

Si comprende agevolmente; che prima che si sviluppassero la teoria e la tecnica della statistica e soprattutto prima che sorgesse la statistica ufficiale, in mancanza di meglio, gli studiosi ricorressero alla *statistica congetturale* e alla *rilevazione indiretta* e mediata e che i pochi elementi o i frammenti di un fenomeno, che si possedevano, si cercasse d'integrare per mezzo del calcolo, per via d'ipotesi, di analogie, di proporzioni, di probabilità.

Della statistica congetturale si servì largamente la scuola degli *aritmatici politici*. Graunt e William Petty l'adoperarono nello studio della mortalità di Dublino e di Londra, di Parigi e di altre città nel 1681, 1682 e 1683; e il secondo, con intenti politici e più sistematicamente, nella *Political Aritmethic* (1691).

Necker, nel 1784, dal numero delle nascite volle dedurre il numero degli abitanti in Francia, adottando la proporzione di 1 a 25,75; Arturo Young, il celebre agronomo-economista inglese, ripeté i calcoli ed i metodi di Vauban, deducendo la produzione agricola e le condizioni della terra in Francia dalla conoscenza diretta di alcune località.

23. *Metodi della statistica congetturale.* — Coi progressi della statistica non si sono eliminati i ricorsi alla rilevazione indiretta, ma si sono meglio disciplinati e si indicano con maggiore esattezza i vari procedimenti della statistica congetturale.

La rilevazione indiretta si fa per *estimi* e per *proporzioni*,

a) *L'estimo* o *stima* è un apprezzamento in forma quantitativa di un determinato fatto, cui si ricorre quando non può farsi la rilevazione di tutte o di una parte sufficiente delle sue manifestazioni. Nell'*estimo* si può procedere per *approssimazione* o per *analogia*.

1.° *Approssimazione.* Si assumono informazioni da coloro che sono esperti in una determinata materia o che, per ra-

gioni di ufficio. devono conoscere un determinato fatto nelle sue manifestazioni o che, per interesse proprio, sono obbligati a tener dietro alle suddette manifestazioni del fenomeno. Esempio: si fa la statistica della produzione agraria per approssimazione, chiedendo notizie ai sindaci, ai comizi agrari, ai grandi proprietari e fittaioli, ai periti agronomi, ecc.

Il metodo dell'approssimazione—fondato sull'apprezzamento personale degli osservatori e quindi ben lungi dalla obbiettività propria del metodo statistico—dà dei risultati, che sono sempre di larga approssimazione, ma che meritano maggiore o minore fiducia, secondo che l'apprezzamento si fonda su una base più o meno sicura di dati positivi, e secondo che è diretto da uniformità ed omogeneità di criteri (*Bosco*).

Basterebbe la statistica della produzione agraria, la cui importanza politica ed economica cresce ogni giorno e richiama l'attenzione degli statistici, degli uomini di Stato, dei commercianti, ecc., per mostrare che della Statistica congetturale non può farsi a meno e che c'è tutto l'interesse politico e sociale nell'organizzarla meglio e nel perfezionarla.

2.^o *Analogia*. Si studiano le condizioni quantitative di un fatto e dalle medesime si deducono approssimativamente le condizioni quantitative di un altro fatto, che ha con quello una relazione di coesistenza e di successione. Così, dal numero delle case, in altri tempi, s'induceva il numero delle famiglie e degli abitanti. Questa era la base del censimento veneto. Giovanni Villani cercò indurre il numero degli abitanti di Firenze dal consumo del pane.

In questa forma la statistica congetturale si adopera sempre meno. Ma è una specie di induzione per analogia quella che si fa calcolando certi consumi dalla quantità di prodotti importati ed esportati.

b) Nella statistica per *proporzionalità*, si può procedere dalla *parte al tutto* o dal *tutto di un fenomeno alla totalità di un altro*.

Questa forma di statistica congetturale si basa sempre sulla *rilevazione diretta* o di una parte del fenomeno o della totalità del fenomeno, da cui s'induce la misura dell'altro fenomeno che si vuole studiare.

1.^o *Dalla parte al tutto*.—Si rilevano le condizioni quantitative di una parte del fatto e si deducono da queste le condizioni quantitative dell'intero fatto. *Esempio*: La statistica postale (movimento delle lettere, giornali, ecc.) si fa per una settimana di media importanza: e dai risultati si calcola per l'intero anno, moltiplicandoli per cinquantadue.

Furono rilevazioni indirette per *proporzionalità* quelle di Vauban, di Young, ecc. poco prima citate.

2.º) *Dal tutto di un fenomeno alla totalità di un altro.*— Si rilevano i dati di un fatto, di un fenomeno, che si considera come causa o come avente influenza sopra un altro fenomeno e s'inducono le condizioni del fenomeno, che si considera effetto o che ha col primo delle relazioni; o viceversa, si prendono i dati del fenomeno effetto e si risale alle condizioni quantitative del fenomeno causa. *Esempi*: Dal prodotto totale di certe imposte s'induce la ricchezza totale di un paese. Dal numero dei reati, delle prostitute, dei figli illegittimi, ecc., si argomentano le condizioni morali di un paese: si fa la cosiddetta *statistica morale*.

Questa statistica si chiama anche *indiziaria*, perchè ordinariamente si prendono alcuni fenomeni sintomatici, caratteristici, per darli come *indici*, come esponenti sintetici, di un aspetto generale della vita sociale.

Alcuni hanno erroneamente ridotto alla *statistica indiziaria* tutta la *statistica congetturale*, che, come abbiamo visto, comprende diversi processi per valutare le manifestazioni statistiche.

Per quante critiche si possano rivolgere alla *rilevazione indiretta*, non è facile negare ai suoi metodi il carattere statistico; ch'è massimo, ad esempio, nella *statistica morale*. Nella quale la rilevazione dei dati, che si riferiscono ai reati, agli illegittimi, ai suicidi, ecc., in realtà è *diretta* e non è *congetturale*: l'induzione, la congettura cominciano nell'apprezzamento del dato.

Le cosiddette *investigazioni rappresentative* del Kiaer in fondo non sono che una applicazione del metodo della *Statistica congetturale*, che induce dalla *parte al tutto*.

In Norvegia si sono fatte già con successo delle investigazioni rappresentative sulle condizioni della popolazione e sulle professioni nei comuni rurali e urbani, per determinare il reddito, ecc. Si procedette a siffatta investigazione circoscrivendola ad un numero relativamente ristretto di persone, ma assortite in modo che la composizione di codesto piccolo gruppo di popolazione stesse a rappresentare la composizione, il tipo della intera popolazione.

Si procedette via via per successive limitazioni, si scelsero un certo numero di città—23 sulle 61 esistenti nel regno — e comuni rurali = 128 su 498 — in modo che tra le città fossero rappresentate così le grandi, come le medie e le piccole, e che nei comuni rurali vi si trovassero i vari tipi di

agricoltura e di industrie agrarie prevalenti nel Regno. Le persone, che furono oggetto dell'indagine, per ridurre il numero delle osservazioni, furono di determinare età scelte come caratteristiche di 5 in 5 anni. I risultati di queste investigazioni furono controllati e trovati molto approssimativi al vero con altri metodi di rilevazione.

CAP. III.

INCHIESTE E MONOGRAFIE DI FAMIGLIA.

24. *Carattere statistico delle Inchieste e delle Monografie di famiglia.* — A complemento alcune volte e tal'altra in sostituzione della rilevazione diretta o indiretta nello studio dei fenomeni sociali hanno assunto una grande importanza le *Inchieste* e le *Monografie di famiglia*, che alcuni statistici eminenti credono che non facciano parte del vero metodo statistico.

Le *monografie di famiglia*, in quanto dalla conoscenza delle condizioni di una o più famiglie addette ad una data professione s'induce alla condizione del resto delle famiglie, che pur presentano tante numerose differenze tra loro, rappresentano sicuramente un metodo congetturale, che si può riattaccare a quelli precedentemente esposti.

È innegabile, però, che tanto le *inchieste*, quanto le *monografie di famiglia* assumono sempre più carattere statistico e sono illustrate da tavole, diagrammi, cartogrammi, numerosi. Rappresentano esempi spiccati di *rilevazione riflessa*, perchè vanno in cerca del fenomeno, che studiano.

23. *Inchieste.* — Le *inchieste* presero il nome da *inchiesta*, che nei dizionari viene presa come sinonimo d'*investigazione* e corrispondendo all'*enquête* dei Francesi, all'*inquiry* degli Inglesi. Mirano ad ottenere notizie o dettagli esatti su particolari fenomeni politici e sociali, I vari aspetti della questione sociale si studiano di preferenza colle *inchieste*.

Le *inchieste* possono essere *private* e *pubbliche*.

Si comprende che l'*inchiesta privata* non può condurre a risultati esatti e completi. Talora mancano i mezzi a coloro che la intraprendono; e coloro che devono somministrare notizie spesso vi si rifiutano, I privati non hanno alcun mezzo per costringerli a darle, molto meno possono invocarne la punizione se le danno incomplete o false.

Alcune *inchieste*, sebbene intraprese da singoli individui,

hanno un carattere pubblico, perchè eseguite per incarico e colla cooperazione di un ufficio pubblico*.

È evidente che, quando l'*inchiesta* concerne un fenomeno, che per la sua vastità e pel suo interesse generale non può essere condotta bene dai privati, dev'essere pubblica.

Le *inchieste pubbliche*—politiche, amministrative, sociali—possono essere *governative e parlamentari*. Nelle prime, non c'è l'obbligo nei privati di ottemperare alle richieste di coloro che le eseguiscano. Ma l'*inchiesta* se anche ordinata da uno dei due rami del Parlamento, se non ha tutti i caratteri dell'*Inchiesta parlamentare*—i caratteri di una vera legge—non impone agli interrogati il dovere di rispondere**.

Si ordinano le *inchieste pubbliche*, governative o parlamentari, quando qualche grave avvenimento ha richiamato l'attenzione del pubblico, del Governo e del Parlamento; e quando si vogliono fare alcune leggi importanti, per conoscere tutte le condizioni che si riferiscono alla materia †.

* Appartiene a questo tipo d'*inchieste* quella del Jenks sui *Trust* negli Stati Uniti del Nord-America. Essa, però, non condusse ai risultati sperati perchè la maggior parte dei *Trust* si rifiutarono recisamente a dare le notizie sui salari, sui prezzi dei generi prodotti sui profitti ecc., ecc.

** I Governi preferiscono le *inchieste* governative e amministrative i cui comitati inquirenti vengono nominati per Decreto reale e i cui poteri vengono pure designati nella stessa guisa. In queste *inchieste* c'è sempre modo di limitare le domande, evitando quelle incommode, che possono fare i troppo esigenti membri del Parlamento e si può anche nascondere in tutto o in parte i risultati—a meno che il Governo non sia impegnato formalmente a pubblicarle nella loro integrità. L'*inchiesta* governativa sulle Banche di emissione nel 1889 riuscì a scoprire il marcio che c'era nella Banca Romana; ma il ministro del tempo non credette conveniente farlo noto—anzi il ministro di agricoltura, on. Miceli, dette alla Camera un riassunto della famosa relazione Biagini del tutto falsa. L'*inchiesta governativa* sulla stessa quistione eseguita nel 1893 sotto il Ministero Giolitti, siccome era state promesso alla Camera che se ne sarebbero pubblicate le risultanze, condusse, alla *relazione Finale*, che conteneva buona parte della verità.

† In Italia sono celebri le *inchieste parlamentari* sul brigantaggio (1863-64): due se ne fecero sulle condizioni politiche e sociali della Sicilia (1867 e 1875): corre per la bocca di tutti, pur essendo poco conosciuta, e del resto poco meritevole di esserlo, l'*inchiesta agraria*. Si fecero *inchieste* o *governative* o *parlamentari* sulle opere pie, sulle ferrovie, sulle condizioni igieniche dei lavoratori della terra ecc. Se n'è fatta un'altra *testè* sulle condizioni dei contadini nel mezzogiorno e in Sicilia.

La maggior parte delle *inchieste* italiane non sono state seguite dai provvedimenti che vennero indicati come opportuni o indispensabili; d'onde l'incredulità del pubblico sulla utilità delle *inchieste* che anzi sembrano destinate ad ingannare o deviare la pubblica opinione.

26. *Metodi per le inchieste.* — Si procede alle *inchieste pubbliche* o colla interrogazione *diretta* degli interessati fatta dai membri delle relative *Commissioni*: o *indirettamente*, cioè collo invio per mezzo della posta e con altro mezzo dei *quistionari*. I due metodi non si escludono e possono anzi completarsi a vicenda. Certamente è più facile e più economico il secondo e lo si adotta più di frequente nelle *Inchieste* economiche e industriali molto vaste.

Gl'inconvenienti del secondo metodo sono gravi. Talora le domande del quistionario non sono ben chiare e non vengono esattamente comprese da coloro, cui sono dirette: le risposte non possono essere, quindi, che incomplete o inesatte. Talaltra, è facile nascondere la verità, da chi ha interesse a nascondersela; così in un'inchiesta sui salari: sul lavoro notturno delle donne e dei fanciulli, sulle condizioni igieniche o degli opifici; ecc. l'industriale, il capo dell'opificio ha tutto l'interesse a nascondere la verità, rispondendo esso per gli operai. Invece, l'inchiesta condotta sul luogo, direttamente, dai membri della Commissione o da speciali funzionari—ne hanno gli *Uffici del lavoro* in America, in Inghilterra, nel Belgio ecc.—riesce a scoprire la verità apprendendola dalla bocca degli operai, che hanno spesso un interesse antagonistico a quello degli industriali. Si aggiunga, inoltre, che sul luogo l'osservazione *de visu* delle condizioni reali può suggerire a chi fa l'inchiesta qualche domanda interessante, che pur non sia compresa nel *quistionario*.

Ad ogni modo, quando si procede per invio di *quistionari* alle persone, che devono rispondere e somministrare le risposte, bisogna mettere dell'attenzione nel formulare le domande o i *quesiti*.

I *quesiti* dice il Gabaglio, devono essere:

a) *Sobrii*, cioè restringersi ai dati più necessari e che si ha fiducia di potere ottenere. Non si deve chiedere molto perchè pur chiedendo poco, molto si deve faticare per ottenerlo; nè si devono fare domande insidiose, se non si vuole far nascere il sospetto, che principale movente della rilevazione sia l'interesse fiscale.

b) *Chiari*, affinché siano facilmente compresi da tutti coloro a cui sono indirizzati. Devono, perciò, essere espressi nel linguaggio più familiare e colle parole più comunemente usate.

c) *Specifici*, tali cioè che non possano essere variamente interpretati. So si prevede che qualche dubbio possa sorgere

1e nel modo d'intenderli si deve toglierlo addirittura con una parola, con una nota, con una spiegazione opportuna.

d) *Semplici*, cioè chiedere risposte brevi, che possano esser date anche dai meno istruiti.

d) *Coordinati*, cioè legati fra loro in modo che si possa per mezzo di essi, sindacare le risposte. Volendosi, per esempio, conoscere l'estensione delle varie specie di terreni produttivi in ciascun comune, non basta domandare la superficie delle terre arabili, dei prati, dei boschi, dei pascoli, dei vigneti, ma anche quella dei terreni improduttivi, e l'estensione totale dell'una e dell'altra, servendo quest'ultimo dato come mezzo per verificare gli altri e potendo esso medesimo essere riscontrato colle indicazioni del catasto.

f) *Categorici*, cioè tali che costringano a dare risposte determinate, non vaghe. La risposta deve consistere o in una cifra numerica (Età? 20 anni), o in una brevissima indicazione (Professione? *Mugnaio*), o in una affermazione o in una negazione (Esistono filande in codesto comune? *Si o no*); non mai in una indicazione indeterminata, quale sarebbe: *poco, molto, sovente, di rado, grande, piccolo* *.

27. *Monografia di famiglia*. — L'*inchiesta*, come si è visto, può essere *pubblica* o *privata*. La ricerca per mezzo delle *monografie di famiglia* è quasi sempre privata o intrapresa a cura di qualche associazione, come in Francia, dove molte se ne devono alla *Société d'Économie Sociale*, che s'ispira alle dottrine economiche e religiose del Le Play. Questi ne fu l'inventore e ne pubblicò la collezione nelle sue opere: *Les ouvriers Européens*. L'*inchiesta* dice Cheysson, è la statistica ufficiale con tutti i suoi pregi: la monografia è l'opera per eccellenza dei privati. L'*inchiesta* guarda alla superficie; la *monografia* alla profondità; l'*inchiesta* e la statistica ufficiale si curano della quantità; la *monografia* mira soprattutto alla qualità dei fenomeni sociali osservati.

Monografia è per i lessici un lavoro, che tratta un punto particolare di letteratura e di scienza; ma anche quà l'uso l'ha applicata a significare: "l'investigazione profonda di un caso determinato, appartenente ai fenomeni sociali, opportu-

* *L'Ufficio del lavoro* di Milano istituito dalla *Società Umanitaria Loria* procede alle *Inchieste* sul lavoro e sulle condizioni dei lavoratori, distribuendo degli eccellenti *Questionari* e mandando sul luogo degli incaricati speciali, che arrivano preceduti quasi sempre dai primi. I *questionari* sono assai dettagliati e opportunamente sono differenziati secondo che servono pei funzionari dello Stato e per le autorità locali, pei lavoratori o pei proprietari o industriali.

namamente trascelto come tipo di altri analoghi dalla personale esperienza dell'osservatore „ (Salvioni). La *Monografia*, dice Cheysson è lo studio approfondito di un *soggetto* scelto come *tipo*: un individuo, una famiglia, un opificio, un comune, una nazione „

Donde la denominazione di *Metodo Tipologico* dato a questa forma di ricerca,

I caratteri specifici del metodo monografico in generale consistono: a) nel limitarsi allo studio d'una istituzione o di un fenomeno sociale nettamente distinto da altri; b) nel descriverlo minutamente nei suoi particolari e nelle sue attinenze in un dato momento e luogo pur così bene delineati; c) nel seguire non l'ordine od il disegno, che meglio piaccia all'osservatore; bensì uno schema metodicamente prestabilito, quasi una trama su cui man mano l'osservatore intessa l'ordito; d) nell'essere quindi comparabile con altre monografie fatte sullo stesso schema, in guisa che risultino chiare le differenze o le analogie, che il fenomeno o l'istituto esaminati presentano col variare dei luoghi, dei tempi, dell'ambiente sociale „ (Bosco).

Le *monografie di famiglia* riguardano a preferenza quelle operaie, perchè queste ultime sono le più numerose e sono le più importanti: danno a tutte le società i mezzi di sussistenza. Questo studio ci fa conoscere il disotto delle società. La famiglia operaia, inoltre, è più esposta all'influenza della costituzione sociale, non avendo le risorse e le abitudini delle classi più elevate per sottrarvisi; essa fa apparire meglio delle altre, per lo stesso motivo, il carattere etnografico di una società, specialmente se è stabile. E Le Play non studiava che le famiglie, che egli considerava stabili.

La monografia di famiglia serve a correggere la facilità delle generalizzazioni, tanto dannosa nelle scienze sociali. Nello studio delle classi lavoratrici, permette di ricercare le cause delle loro condizioni, i rimedi e i risultati; ne fa constatare le abitudini, il malessere, le sofferenze, i difetti, le virtù, i sentimenti, i pregiudizii, le aspirazioni; ne mette in evidenza al vivo tutti i fattori tecnici e morali, che influiscono sulla famiglia, la sua classificazione, il suo stato di pace o di antagonismo coll'ambiente sociale.

Ma la *monografia di famiglia*, osserva il Salvioni, imita il metodo di ricerca nelle scienze naturali colle osservazioni singole, e si mette in contrasto colla statistica, che si caratterizza dalla rilevazione *in massa*; ricorre al metodo deduttivo

prescegliendo *a priori* il tipo che deve studiare; mentre la statistica argomenta *a posteriori* sui risultati raccolti.

Il carattere statistico, però, può rilevarsi ad esempio, da tutte le notizie sul salario dei membri della famiglia studiata, dal bilancio ecc.

8. *Partizione della monografia di famiglia.* — Rinunziando per amore di brevità, ad esporre tutte le precauzioni, che si devono prendere nel procedere alla compilazione di una monografia di famiglia riassumerò ciò che deve formare oggetto dell'osservazione. Ciò servirà a far meglio comprendere l'utilità dello studio.

Ogni monografia si compone di tre parti: a) le osservazioni preliminari che definiscono le condizioni dei diversi membri della famiglia; b) il bilancio delle entrate e delle spese; c) le *note* o elementi diversi della costituzione sociale, consacrate ai fatti generali che s'incontrano nella famiglia e alle conclusioni.

a) *Osservazioni preliminari.* — Vi si devono comprendere queste notizie: stato del suolo, della produzione locale e della popolazione; stato civile della famiglia; religione e abitudini morali; igiene e servizio sanitario; rango della famiglia. Studio sulla proprietà immobiliare e mobiliare; sovvenzioni diverse: usufrutto — ad esempio, le abitazioni che ricevono dagli industriali; — diritto di uso, prestazione di oggetti e di servizi — ad esempio, confezione gratuita dei vestiti, servizio medico gratuito, somministrazione di carbone gratuito, a prezzo di costo ecc., ecc.; lavori e professioni — osservare quali membri della famiglia lavorano, se il salario è in danaro o in prodotti; se il salario è a giornata, a mese, ad anno o a cottimo; ore di lavoro e loro rapporto coll'altezza del salario; lavoro della donna e dei fanciulli e conseguenze igieniche, economiche e morali; modi di esistenza: alimenti e pasti — proporzioni di cereali, sostanze grasse, latte e uova, carne e pesce, legumi e frutti, condimenti e stimolanti, bevande fermentate, alimenti che si portano dalla casa sul luogo del lavoro; abitazioni, mobili e vestiti, riscaldamento, illuminazione, ricreazioni, compresa la coltivazione dei piccoli giardini; storia della famiglia, che comprende le fasi diverse della sua esistenza e lo studio delle istituzioni e dei costumi, che ne assicurano il benessere fisico e morale; società cooperative, assicurazioni diverse.

b) *Il bilancio.* — Bilancio delle entrate, bilancio delle spese, conti annessi. Si connette tutta la seconda parte alla prima, che riassume mettendone in evidenza i risultati.

Percentuale della spesa di una famiglia
Operaia Classe media Agiata

Sussistenza.	62	95	90	85
Vestiti	16			
Alloggio.	12			
Riscaldamento e illuminazione	5			
			55	50
			18	18
			12	12
			5	5
Istruzione	2	5	10	15
Protezione legale	1			
Igiene e salute	1			
Confort mentale e ricreazioni	1			
			3,5	5
			2	3
			2	3
			2,5	3,5

Queste proporzioni sono state confermate da ricerche di altri scrittori e da inchieste ufficiali. Secondo i dati di Engel nelle famiglie, il cui reddito è sotto 600 lire, il solo cibo prende il 71,5 del totale. Questa parte discende gradatamente sino al 64,8 in quelle, che hanno un reddito superiore a L. 2000.

La percentuale dell'alimentazione è più elevata nelle famiglie più povere; diminuisce a misura che migliora la sua condizione economica e crescono le entrate. Viceversa si eleva la percentuale delle altre spese, specialmente quelle del secondo gruppo.

La parte del bilancio consacrato alla spesa serve a misurare le variazioni del valore della moneta.

Insistere soprattutto sulle variazioni dei singoli *capitoli della spesa* in rapporto alle variazioni della *entrata totale*.

c) *Le note*. — Raccoglie le notizie sull'ambiente sociale in cui la famiglia si muove; le cause, che ne hanno cagionato la dissoluzione o mantenuto la stabilità; rileva il regime di trasmissione dei beni; i principali caratteri del regime industriale locale, ecc.

Si avrà un'idea di questa parte dal titolo delle *note* della monografia di una famiglia di un tornitore meccanico dell'officina Cockeril a Seraing (Belgio): storia o organizzazione dello stabilimento Cockeril; inchiesta ufficiale sulle condizioni del lavoro nel Belgio; società di mutuo soccorso *I meccanici riuniti*; mantenimento della pace sociale nelle società anonime industriali e patronato; industria vetriera e insufficienza della sola elevazione del salario per migliorare il benessere dei lavoratori; manifestazione dello spirito di associazione negli operai belgi; accesso degli operai alla proprietà del focolare domestico; situazione economica e sociale del Belgio.

Oggi sono gli Stati per mezzo degli Uffici del lavoro, che fanno fare delle monografie di famiglia, molto adatte a far conoscere le condizioni reali delle famiglie operaie.

29. *Altre forme della monografia.* — La *monografia di famiglia* è lo strumento di ricerca più noto e finora più usato per lo studio della vita o della economia delle classi lavoratrici.

Nella pratica e nella tecnologia statistica si hanno altri strumenti di rivelazione statistica, che completano ed allargano il campo d'investigazione della monografia di famiglia.

a) *Monografia di opificio.* Ne ha compilato il quadro M. E. Cheysson fin dal 1887 *. La monografia di opificio, secondo l'Autore, si compone di due parti principali: l'una riguardante la organizzazione commerciale; l'altra l'organizzazione del lavoro.

1.° *Organizzazione commerciale.* Generalità sulla regione. Generalità sulla industria. Organizzazione finanziaria. Organizzazione commerciale.

2.° *Divisioni principali dell'opificio.* Reclutamento e distribuzione del personale. Salarii. Durata del lavoro. Istituzioni create in favore dei lavoratori. Abitudini generali della famiglia operaia. Rapporti tra il capitale ed il lavoro.

b) *Monografia di classe professionale.* La monografia di famiglia ha un campo d'investigazione ristrettissimo nel tempo e nello spazio, e non riesce a rendere l'immagine esatta della classe professionale della famiglia studiata, quando in quella son compresi diversi gruppi e sottogruppi di lavoratori ed è vasto il territorio sul quale la classe si estende o vive.

Si può supplire al difetto o aggiungendo note e appendici alla monografia, o compilando parecchie monografie di famiglia. Nell' un caso e nell' altro si fa lavoro incompleto e si spezza l' organicità di trattazione dell' argomento. E miglior processo tecnico per una larga e profonda investigazione di gruppi di lavoratori simili, la *monografia di classe professionale*. In questa, la monografia di famiglia assume il posto di un capitolo fondamentale della ricerca; le *osservazioni sugli elementi diversi della costituzione sociale* assumono il grado di parte principale del lavoro; e così la ricerca analitica e la sintetica si controllano a vicenda, e si completano in uno studio organico e relativamente breve.

La monografia di classe è ancora da costituire nel metodo

* Cheysson: *Bulletin de l'Institut international de statistique*. Tome II, 1887:

e da tradurre in pratica. In Italia ne ha fatto un eccellente tentativo il D.r Baglio: *Ricerche sul lavoro e sui lavoratori di Sicilia. — Il Solfarajo.*

c) *Monografia di professione.* Di fronte alle industrie estese e con grande divisione di lavoro, la monografia di opificio presenta le stesse manchevolezze della monografia di famiglia di fronte alle grandi ed estese classi professionali. E' perciò sorto il voto degli studiosi, che si costituisca la *monografia di professione.* Questa dovrebbe investire tutta l'opera dei soggetti economici direttamente e prevalentemente interessati in una industria, designarne la figura giuridica e sociale, esaminarne l'attività politica e morale.

d) *Monografia di comune.* La serie degli strumenti di rilevazione monografica finora progettata si completa con la *monografia di comune.*

Questa volge su tutta la vita fisica, economica, morale, giuridica e politica d'un grande agglomerato di uomini, e costituisce, o dovrebbe costituire, la sintesi organica delle monografie anzidette. Di essa lo Cheysson ha compilato il quadro per i comuni agricoli, dividendolo in 2 parti.

Parte I. INTRODUZIONE STORICA: 1.º Storia demografica. 2.º Storia economica. 3.º Storia sociale.

Parte II. SITUAZIONE ATTUALE DEL COMUNE: 1.º Descrizione fisica. 2.º La popolazione. 3.º Emigrazione ed immigrazione. 4.º Divisione della proprietà. 5.º Contratti agrari. 6.º Culture. 7.º Istruzione agricola. 8.º Industrie rurali. 9.º Salari e mano d'opera. 10.º Condizioni del personale agricolo. 11.º Risultati economici. 12.º Sindacati agricoli. 13.º Previdenza. 14. Assistenza. 15.º Stato morale e sociale del comune.

Seguendo le linee tracciate dallo Cheysson, allo studioso di economia e di statistica non riesce difficile formarsi il quadro di monografia di un comune agricolo. Stimiamo opportuno avvertire, che nel quadro dello Cheysson si dà sviluppo eccessivo alla storia della popolazione, mentre basterebbero in proposito pochi e rapidi cenni.

Le cinque monografie, delle quali abbiamo parlato, costituiscono una serie completa e continua, poichè si sale con esse a grado a grado dall'individuo all'ente giuridico e collettivo, dall'economia individuale alla sociale, dal lavoro casalingo a quello de' grandi opifici, dall'infimo al sommo della scala sociale.

CAPO IV.

ERRORI DI RILEVAZIONE E CRITICA DEI DATI.

30. *Errori di rilevazione.* — Nella rilevazione dei dati statistici si può incorrere in errori, come in ogni altro studio. Gli errori nella rilevazione hanno grande importanza, perchè ad operazione compiuta assai difficilmente lo studioso, che prende in esame i risultati può accorgersene e nella massima buona fede continuerà ne' calcoli e nelle induzioni, basandosi su dati inesatti. D'onde la necessità di schivarli.

Gli errori nella rilevazione derivano da *cause oggettive* o *sogettive*.

a) Sono *oggettive* le cause, quando sono insite, connaturate col fatto, col fenomeno rilevato, ed hanno origine: dalla *sede*, dalla *qualità*, dalla *complicazione* dello stesso fenomeno.

1.° *Sede*: quando il fatto osservato producesi contemporaneamente in vastissima sede. D'onde, ad esempio, la difficoltà di fare una esatta statistica della *produzione agraria*.

2.° *Qualità*: quando il fenomeno è variatissimo nella sua manifestazione. Ad esempio: la *statistica dei prezzi*.

3.° *Complicazione*: quando il fenomeno si può studiare sotto vari aspetti e si complica o suddivide ognora di più. Così, la crescente specializzazione delle arti, delle professioni, rende spesso inesatte le classificazioni. Perciò, in Germania, nella *statistica delle professioni* si hanno 153 tipi, con 6000 modi di applicazione. Nella sola medicina e chirurgia si distinguono gli ostetrici, gli oculisti, i pediatri, i laringologi, i dermatologi, ecc.

b) Gli errori da cause *sogettive* si riferiscono all'osservazione e derivano da *ignoranza* o da *mala fede* nelle persone, che raccolgono i dati o che li denunciano. Si spiega benissimo l'errore da ignoranza. Ad esempio, molte persone avanzate negli anni ed incolte non sanno indicare la propria età. Le donne spesso non vogliono dirla esatta di proposito. Da ignoranza o da mala fede derivano spesso le erronee dichiarazioni degli emigranti, che non sanno o non vogliono dire se lasciano il proprio paese *permanentemente* o *temporaneamente*. D'onde i salti, che talora sembrano inverosimili, tra l'*emigrazione temporanea* e la *permanente* in talune provincie o regioni.

In quanto alla *mala fede* si ricordi, che taluni impiegati,

per risparmiare lavoro, inventano le statistiche; tal'altra le gonfiano, ecc.

Dice il Messedaglia, che in coloro che fanno le statistiche occorre: *scienza, arte, onestà*. Occorre pure qualche qualità negativa, ch'è tra le più difficili a riscontrarsi: *l'assenza di passione*.

Si ritiene ordinariamente, che gli errori derivati da ignoranza, da pregiudizi e da mala fede non siano correggibili e non si possono scoprire; e questa credenza contribuisce a generare la categoria degli *scettici della statistica*.

La credenza è infondata o esagerata. Non mancano modi per controllare l'esattezza, per lo più approssimativa — su questo è bene intendersi — dei dati rilevati ed elaborati. Intanto è bene ricordare, che la leggerezza e la mala fede dei raccoglitori si scoprono talora da loro stesse.

Gli errori di rilevazione, infine, secondo la causa che li produce, si dividono in *costanti, variabili ed accidentali*.

« Gli *errori costanti*, in quanto sono prodotti da una determinata causa, agiscono in una stessa direzione. Il loro effetto si viene così accumulando a misura che le osservazioni, permanendo immutate le circostanze e le cause di errore, si ripetono. Questi sono i *blasted errors* degli inglesi „ (*Bosco*).

Il censimento può fornire l'esempio di questi tre diversi errori.

Si ha l'errore da *causa permanente o costante*, se si fa il censimento sempre in un giorno, in cui c'è un eccezionale movimento della popolazione; da *causa variabile*, se al censimento si procede ora in un giorno, ora in un altro; da *causa accidentale*, se il giorno del censimento è eccezionalmente piovoso o perturbato da un avvenimento straordinario: una rivoluzione, ecc.

Gli errori non sono sempre eliminabili, ma si deve tentare di evitarli. A ciò in qualche modo si riesce seguendo queste regole:

a) Si modificano le *condizioni estrinseche* della rilevazione; ad esempio: alla rilevazione *riflessa*, quando è possibile, si sostituisce quella *automatica*.

b) Si modificano le *condizioni di tempo*: alla rivelazione *periodica, occasionale* si sostituisce quella *continua*.

c) Si modificano gli *organi raccoglitori*. Così, le statistiche del lavoro fatte dalle Prefetture si sono trovate inesatte per inesperienza del personale e per interesse politico di chi presiede — ogni prefetto vuol far comprendere che sono ottime le condizioni del lavoro nella propria provincia! — ;

perciò dappertutto si fanno fare da appositi *uffici di statistica del lavoro*.

d) Si rinuncia alla rigorosa precisione e si danno i risultati come approssimativi: così, con particolarità nella statistica dei *prezzi*, dei *salari*, ecc.

31. *Critica dei dati*. — I dati statistici rilevati, spogliati ed aggruppati, si devono sottoporre a revisione; se ne deve fare la critica per assegnare ad essi il grado di fede, che possono meritare in tutti gli stadi successivi della loro elaborazione e nel loro uso.

Quetelet distingue la *critica in esterna ed interna*, ovvero *materiale e morale*.

La *critica esterna* si riferisce alla precisione *materiale* dei dati e alle cause estrinseche, che possono influire sul valore del materiale statistico e sono svariatissime, come si sa dagli *errori di rilevazione*. Nella *critica esterna*, si deve badare soprattutto all'esattezza delle somme, che si devono fare con tutti gli avvedimenti possibili. Si cercherà di scoprire e correggere le eventuali contraddizioni fra il dato generale e la somma dei dati particolari.

La *critica interna o morale* è assai più interessante. Mira ad accertare la precisione intrinseca ossia l'esattezza e l'integrità dei dati: si deve esaminare, non solo se il dato corrisponde alla manifestazione *quantitativa* del fenomeno, ma anche se tra i dati se ne comprendano altri di *qualità* diversa cioè non omogenei. La *critica interna* è quella che ha di mira soprattutto gli errori da *cause oggettive*.

Che nel dato vi sia un errore e non sia degno di fede, spesso ce ne possiamo accorgere o possiamo sospettarlo dalla considerevole differenza quantitativa coi dati di altre regioni in condizionianaloghe o con quelle degli anni antecedenti o susseguenti.

Per mezzo di buoni confronti si assegna il grado di credibilità, che si deve accordare al dato. Si procede alla correzione, nei limiti del possibile: comparando i dati attinti a fonti diverse: comparando i dati di un anno con quelli dei dati degli anni antecedenti e susseguenti; paragonando il dato ottenuto dagli organi raccoglitori con altri dati, coi quali è in correlazione; sostituendo al dato originario uno *medio* o *perequato* quando non è stata possibile la correzione. E come si procede può rilevarsi da diversi esempi.

a) Si confrontano i dati in esame con quelli della stessa specie ottenuti da precedenti rilevazioni e debitamente accertati. Così, col censimento del 1881, in Italia il sesso ma-

schile risultò superiore a quello femminile, mentre nella maggior parte dell'Europa centrale e settentrionale prevaleva il sesso femminile. Il dubbio sorgente da questo contrasto venne eliminato dalla circostanza, che anche nei due precedenti censimenti prevalse in Italia il sesso maschile. Di più, dai dati sulla mortalità per sesso si vide, che la mortalità maschile non eccedeva quella femminile tanto, quanto era la eccedenza nella natalità dello stesso sesso. Anzi secondo la proporzione tra la natalità e la mortalità dei maschi, avrebbe dovuto trovarsi nel censimento del 1881 una maggiore proporzione di maschi; e se non si trovò quale doveva essere, anche tenendo conto dell'emigrazione distinta per sesso, ciò si attribuì alla parte dell'emigrazione stessa, che avviene clandestinamente per vari motivi, per sfuggire agli obblighi di leva, perchè non si vuole accordare il passaporto, ecc.

b) I dati si mettono a confronto con quelli della stessa specie, attinti a diverse sorgenti e tutte autorevoli. La concordanza con queste o la piccola divergenza farebbe presumere la loro esattezza. Esempio: mettendo a confronto il numero degli analfabeti per 100 maschi a 20 anni, secondo il censimento del 1881, con quello degli analfabeti per 100 coscritti della leva dei nati del 1860, che costituivano la categoria dei giovani censiti, si trovò che i risultati delle due rilevazioni erano abbastanza concordi.

c) Il controllo si può fare con una esatta operazione di calcolo. Esempio: si vuol vedere se il risultato del censimento della popolazione del 1881 fu esatto. Ai 26,801,154 abitanti del censimento del 1871 si aggiungono i 10,221,316 nati nel 1872-81 e gl'immigrati; si deducono gli 8,200,526 morti e i 200 mila circa emigrati; si trova che il censimento del 1881 dava una popolazione di 28,459,626, mentre avrebbe dovuta darla di 28,561,943. La differenza in meno in una così grossa cifra è di 100,000 circa ed è incalcolabile: appena di 0,3 ‰. In gran parte si ritiene dovuta all'emigrazione.

La differenza tra la popolazione calcolata e quella censita nel 1901 fu maggiore di quella osservata nel 1881; di oltre 300 mila cioè di circa l'1^o/₁₀. E ciò per l'aumentata emigrazione. Fu molto maggiore la differenza tra la popolazione *legale* (35,845,048) e la popolazione *presente* o di *fatto* (34,671,377) nel censimento del 1911.

d) Si cerca di attenuare l'errore del dato, che sembra non rispondente alla verità, ricorrendo alle cosiddette *perequazioni*, cioè sostituendo a quello originario un dato *medio* che risulta dalla somma del primo coi dati di parecchi anni pre-

cedenti e susseguenti e il cui prodotto si divide pel numero dei termini, come si apprenderà dalla trattazione dei *valori medi*. In queste caso, contro il meritato rimprovero che si fa all'uso delle *medie*, cioè di sostituire alla cifra reale una ipotetica, il valore *perequato* serve a sostituire a quello originario uno, che maggiormente si approssima alla realtà.

La *critica* dei dati induce il Gabaglio a queste conclusioni: 1.° La verificazione dei dati statistici primitivi è più o meno facile secondo la natura dei fatti cui si riferiscono. Per alcuni, l'accertamento quantitativo può essere fatto con rigore, ad esempio, pei dati, che si ottengono dai registri *dello stato civile*; altri sfuggono ad una verificazione rigorosa. 2.° Gli errori di rilevazione, di spoglio e di aggruppamento non possono essere eliminati interamente nemmeno coi metodi di rilevazione più perfetti. Ma le varie correzioni ci danno una maggiore approssimazione alla verità.

B) Elaborazione.

CAP. V.

A) OPERAZIONI TECNICHE E LOGICHE.

32. *Spoglio e aggruppamento*. — I dati in qualunque modo siano stati rilevati o raccolti, debbono essere elaborati.

L'*elaborazione* dei dati statistici è una delle parti importanti della teoria della statistica. Vi si procede per renderli comparabili più facilmente, per rendere più chiara ed immediata la comprensione dei rapporti di grandezza, che essi hanno con altre quantità statistiche, e per renderne più agevole lo studio: specialmente per favorire la *interpretazione*, cioè la scoperta dei rapporti causali e delle leggi empiriche.

Le operazioni tecniche e logiche dell'*elaborazione* sono lo *spoglio*, l'*aggruppamento* e la formazione delle *serie*.

La prima operazione tecnica è quella di *spogliare* i dati per poterli *aggruppare*.

Lo *spoglio* si fa rilevando ad uno ad uno i dati individuali dalle *schede individuali*, dai *registri*, dai *questionari*, dalle *schede di famiglia*, e trascrivendoli sopra altrettante asticelle, schede, o cartoline, che si dicono *carte di numerazione*.

La *carta di numerazione* deve contenere un solo dato o un solo carattere di quelli rilevati. Così, da una *scheda individuale*, nella quale è segnato il sesso, l'età, lo stato civile, il grado d'istruzione, la professione, ecc. di un individuo, si trasportano separatamente in altrettante carte di numerazione queste singole condizioni.

Le *carte di numerazione* per ogni singola qualità si addizionano — si *aggruppano*, si *totalizzano*, come dice Bertillon.

Alla fine dell'operazione, in 100 schede individuali si è in grado, perciò, di riconoscere, ad esempio, che ci sono 51 maschi e 49 femmine; 40 sotto i 20 anni; 25 da 21 a 40 anni; 15 da 41 a 60 anni; 10 da 61 a 65 anni, ec., ec.; 58 celibi e nubili, 36 coniugati e 6 vedovi; 60 analfabeti, 38 che sanno leggere e scrivere, 2 con istruzione superiore; 50 appartenenti alla professione agricola, 20 alle industrie, 10 ai trasporti, 2 alle professioni liberali, ec., ec. I risultati dell'aggruppamento poi si trascrivono nella corrispondente colonna, che porta la indicazione di ogni singola qualità, di una *tavola* a tale uopo preparata e che avrà tante colonne quante saranno le qualità rilevate.

Spoglio ed *aggruppamento* vanno di conserva. Lo *spoglio* sarebbe inutile senza l'*aggruppamento*, e questo non sarebbe possibile senza lo *spoglio*.

Mercè lo *spoglio*, da modelli di rilevazione, che non permettono l'*aggruppamento*, si trasportano i dati su altri modelli, che lo consentono. Esempio. si notano i fatti man mano che succedono cronologicamente sopra i registri, poi per aggrupparli si trasporta la totalità di essi su *tavole* aventi le *colonne* opportunamente intestate.

Così, nel censimento del 1881, redatta da ogni famiglia la scheda, gli uffici comunali ne trasportarono i dati su altrettanti bollettini individuali, quanti erano i componenti di ciascuna famiglia. Si ottennero 28 milioni e mezzo di bollettini individuali, quanti furono gli abitanti censiti del regno. Questi bollettini si mandarono alla Direzione generale della statistica, dove furono separati secondo il sesso, l'età, lo stato civile, ecc., degli individui. Bastò quindi numerare i bollettini di ciascun gruppo, per sapere il numero dei maschi e delle femmine, dei fanciulli, degli adulti, dei vecchi, dei celibi, delle nubili ecc.

In Italia, nel 1881, per lo *spoglio* si adoperò un contatore meccanico; negli Stati Uniti, nel censimento del 1890, si adoperò la macchina Hollerit.

Ed apparecchi elettrici e macchine di ogni sorta si adoperano dappertutto per le varie operazioni tecniche della statistica; ma una descrizione accurata, quale la dà, ad esempio, il Salvigni nelle sue lezioni, non riesce a farne comprendere il meccanismo. Bisogna vederle funzionare.

L'*aggruppamento* deve farsi alla periferia dagli *organi raccoglitori*; o al centro dagli *organi dirigenti* della rilevazione?

L'aggruppamento in Italia prima veniva fatto dagli organi *raccoglitori*, ma erano inabili e svogliati. Le inesattezze inoltre derivano dalla differenza del metodo, mentre la mancanza di mezzi tecnici rendeva più difficile l'operazione. Preferiscansi per ciò gli organi dirigenti. Gli uffici locali, però, potendo avere bisogno di documenti originari, si può riparare o restituendoli o riproducendoli.

Ciascun gruppo delle unità raccolte, spogliate ed aggruppate o ridotte ad espressione quantitativa, rappresenta un *dato statistico primitivo*. Il quale può definirsi: " *la somma delle manifestazioni di un fenomeno in un dato tempo o in un dato luogo.*"

I prospetti, nei quali si espongono i *dati statistici primitivi* si dicono *tavole statistiche originarie* o *primitive*. La *tavola* sarà *derivata*, quando presenterà i dati, che hanno subito l'elaborazione col secondo gruppo di operazioni. I dati, però, prima di sottoporsi a queste operazioni matematiche, vengono controllati, sottoposti a critica.

33. *Le serie statistiche.* — Non si possono iniziare le operazioni matematiche della *elaborazione*, senza conoscere che cosa sieno le *serie statistiche*.

La *serie statistica* è la riunione di varie *quantità statistiche* o *dati*.

La *serie statistica* ha queste due caratteristiche:

a) risulta dai gruppi omogenei, considerati come altrettante unità; b) contiene un certo numero di quantità.

Non si può precisare *a priori* il numero delle quantità, che deve contenere una *serie*: egli è certo, però, che si ha una *serie* solo quando il numero delle unità, che contiene, è abbastanza rilevante. Quanto maggiore è il numero delle osservazioni nel tempo, tanto più facilmente si può scoprire nei fatti quello, che essi hanno di regolare e di relativamente normale, in mezzo all'accidentale ed anormale. Se il numero delle quantità statistiche è talmente esiguo, che togliendone una sola, verrebbe alterata sensibilmente, la *media* la riunione di tali quantità non costituirà una *serie*.

Le quantità statistiche aggruppate secondo una certa scala di misura costituiscono la *seriazione*. La *serie* consta di un solo elemento: il dato statistico; la *seriazione*, ne comprende due tra loro connessi: il dato statistico e la scala di misura con cui il dato è messo in rapporto.

" Così se si distribuiscono i proprietari di immobili in un paese secondo l'estensione dei terreni, i contribuenti secondo l'ammontare del reddito tassato o dell'imposta, i condannati dalle

varie magistrature secondo la durata della pena, i coscritti secondo la statura, i censiti secondo l'età ecc., formano, non delle *serie*, ma delle *seriazioni* „ (Benini). Ciò che si comprenderà meglio quando si tratterà del *metodo seriale*.

La *legge seriale* è quella secondo la quale nessun fenomeno avviene isolato, ma è sempre l'espressione di una serie di fatti analoghi.

La *serie* può essere, di *fatto*: per esempio, la distribuzione di una popolazione per professioni; di *tempo* (i nati in Italia dal 1872 al 1900); di *luogo* (i nati in Italia nelle varie provincie).

Le *serie* nel *tempo* riguardano periodi di durata più o meno lunga e possono essere *periodiche* e *non periodiche*. “ Periodiche, quelle delle quali si considera lo svolgimento in un ciclo naturale o convenzionale di tempo (come i dodici mesi dell'anno; i sette giorni della settimana; le ventiquattro ore del giorno) e i cui massimi e minimi si ripetono nei successivi cicli ad intervalli abbastanza regolari. Non periodiche, quelle nelle quali si prescinde dal ciclo e si considera la semplice continuità di manifestazione del fatto nel succedersi del tempo „ (Benini).

Le *serie* nello *spazio* o di *luogo* possono comprendere circoscrizioni territoriali più o meno ampie, da quelle di interi Stati alle altre dei compartimenti, delle provincie, dei comuni: del pari, le *serie* di *fatto* possono scendere a divisioni e suddivisioni. Così ad es. dopo avere esposta la serie di professioni principali in Italia, si potrà vedere come si suddividono le professioni liberali — medicina, giurisprudenza, ingegneria, scienze fisiche, insegnamento ecc. — o la professione agricola — proprietari di terreni, fittaiuoli, mezzadri, braccianti, ecc. ecc.

Si devono seguire certe norme nel disporre i termini di una *serie* in guisa da rendere più evidente il fenomeno o il rapporto che si studia.

La *serie statistica nel tempo* prende nomi diversi secondo le modificazioni, che presentano attraverso al tempo i fatti, che la compongono. Si hanno perciò le serie a *carattere statico*, *dinamico misto* e *indeterminato* o *irregolare*.

a) La *serie a carattere statico* è formata da gruppi o unità statistiche, che non tendono né a crescere, né a diminuire nella loro manifestazione quantitativa. Se il numero dei nati maschi in dieci anni mantiene la stessa proporzione colle femmine e non vi sono che piccole oscillazioni ora in più, ora in meno, si ha la *serie a carattere statico*, nella quale le oscillazioni tendono ad una determinata manifestazione.

b) Nella *serie a carattere dinamico*, i fenomeni nella loro manifestazione quantitativa presentano la tendenza—tenendo conto dell'insieme delle loro oscillazioni — o a crescere o a diminuire. I numeri relativi ai suicidi in venti anni di seguito mostrano la tendenza all'aumento del fenomeno; quelli relativi all'omicidio invece manifestano la tendenza alla diminuzione. I dati sul suicidio e sull'omicidio, in un determinato numero di anni, perciò costituiscono due *serie statistiche a carattere dinamico*.

c) Nella *serie a carattere misto*, i fatti nella loro manifestazione quantitativa passano, dopo un certo tempo, dalla condizione statica alla dinamica o viceversa. Esempio: le nascite illegittime in Baviera: le quali, prima rappresentavano una *serie a carattere statico*; poi a *carattere dinamico* con tendenza alla diminuzione ed in ultimo contendenza all'aumento.

d) Nella *serie a carattere indeterminato*, le oscillazioni in più o in meno nelle manifestazioni quantitative del fenomeno sono tali, che non è possibile stabilire la tendenza, perchè si passa rapidamente dalla tendenza all'aumento a quella alla diminuzione, e viceversa. Es. i prezzi di alcuni generi.

34. *Il peso nella serie.*—Le *serie* si disse che sono una riunione di *dati* o *quantità statistiche*; ma il *dato* o *quantità statistica* può risultare da diverse *unità* o *elementi statistici*. Il numero delle unità o delle manifestazioni di un fenomeno che compongono il *dato*, rappresenta il *peso*.

La *serie* può essere quindi a *peso costante* ed a *peso variabile*. Si dice a *peso costante*, quando le *unità* o *elementi statistici*, che formano i *dati* singoli di una *serie*, non variano nel tempo o nello spazio o presentano tra loro delle piccole oscillazioni: sarà a *peso variabile*, quando le variazioni nel numero delle *unità* di ogni singolo *dato* saranno sensibili. Così, ad esempio, si sa che dal 1872 al 1900 in Italia sono nati circa 106 maschi per 100 femmine. Questo *dato derivato* o *proporzionale* risulterebbe da una *serie a peso costante* se le oscillazioni nel numero delle nascite tra un anno e l'altro fossero state esigue—1872, nati 1,028,682; 1873, nati 985,188; 1874, nati 951,658, ecc. ecc.

Se le nascite fossero state di 1,000,000 nel 1.º anno: di 1,400,000 nel 2.º: di 700,000 nel 3.º di 1,800,000 nel 4.º e così di seguito, la *serie* sarebbe a *peso variabile*, per quanto le proporzioni tra le nascite maschili e le femminili fossero state pressochè identiche tra la *serie a peso costante* e la *serie a peso variabile*.

Il tener conto del *peso* ha una grande importanza nella valutazione di molti fenomeni sociali, come si vedrà quando discorreremo delle *medie*.

Il Salvioni chiama *tipiche* le serie i cui dati si dispongono attorno ad un centro, *non tipiche*, quando non presentano tale disposizione.

CAP. VI.

B) OPERAZIONI MATEMATICHE.

I.° Dei valori mediani.

35. *Generalità; significato delle medie.* — L'elaborazione matematica dei dati comprende due gruppi principali di operazioni: a) la formazione dei *valore mediani* b) la riduzione delle cifre assolute in *cifre proporzionali*. Non mi occupo delle *interpolazioni*, per le quali occorrono conoscenze più larghe di matematiche e sono di uso assai raro nelle scienze sociali.

I fenomeni di cui si occupa la statistica essendo *atipici*, e le *unità statistiche* da un anno all'altro, da un luogo all'altro — cioè nel tempo e nello spazio — presentando tra loro delle differenze in più o in meno, si è sentito sempre il bisogno di ridurle ad una espressione sintetica, che possa servire a dare un'idea approssimativa e generale della manifestazione quantitativa del fenomeno. D'onde le *medie*, la cui formazione costituisce parte importante del secondo gruppo delle operazioni della elaborazione, delle *operazioni matematiche*.

E' comune e frequente l'uso, che facciamo delle *medie* nella vita quotidiana, spesso incoscientemente. Esse servono a dare l'idea generica di un gruppo o serie di fenomeni, di un insieme qualsiasi d'individui, perciò le troviamo implicitamente in ogni nome, aggettivo, pensiero, che non si riferisce ad un determinato fenomeno o ad un determinato individuo, ma a gruppi, o serie o insieme d'individui o di fenomeni. Ciò avviene perchè siamo indotti naturalmente alle indicazioni generiche, con cui una intera classe o parte di classe può essere designata.

Ora la statistica per lo appunto non occupandosi del *singolo*, ma delle *masse*, deve necessariamente servirsi delle *medie*. " Perciò un dato medio deve considerarsi come un tipo riassuntivo una manifestazione sociale in massa oppure come un'astrazione numerica senza carattere tipico „ (Mayr).

Quando e come una *media* può rappresentare il *tipo* di un fenomeno emerge chiaro da queste osservazioni del Salvioni. Sarà un *tipo*, quando i fenomeni parziali, se non corrispondono esattamente al risultato medio, gli girano attorno in più o in meno, gli si accostano di continuo positivamente o negativamente. La media delle nascite e delle morti in una data nazione, se dedotta da una *serie statica*, calcolata sopra un periodo di 25 anni, può essere considerata, come un tipo della *natalità* e della *mortalità* di quel paese. È quindi possibile, che nell'anno 26° nascite e morti si avvicinino a quelle medie, che perciò, in un senso diverso da quella dato alla parola del Quetelet, si diranno *tipiche*.

Altre medie statistiche invece non rappresentano un *tipo* di fenomeni ed hanno un'importanza scientifica minore. Così se dai dati sull'età di ogni singolo cittadino risultanti da un censimento noi ricaviamo gli anni vissuti complessivamente dalla popolazione intera e dividiamo la somma dell'età pel numero dei cittadini, ricaviamo la cosiddetta *età media* dei censiti, e sia di 28 anni e 4 mesi, com'è risultata nel censimento del 1901 in Italia. Questa cifra in apparenza è una *media*, come quelle delle nascite e delle morti: ma in sostanza è cosa assai diversa. Avere l'età di 28 anni non è punto raggiungere un *tipo* per le persone di cui è composta una popolazione, anzi non si può nemmeno pensare al fatto, che esista una popolazione di persone, che abbiano 28 anni o presso a poco.

In questo caso la media è una semplice astrazione del calcolo e non ha che un lontano addentellato con le condizioni dei fenomeni reali. Perciò il Salvioni comincia dal distinguere le *medie tipiche* dagli *adeguati numerici*.

La ragione per cui si ricorre alle *medie* da Gabaglio viene indicata nel fatto, "che i dati costituenti una serie non sono mai tutti fra loro perfettamente uguali (se la serie è di carattere statico) nè sempre palesano un movimento regolare (se la serie è di carattere dinamico): imperocchè i fatti, che essi rappresentano, sono bensì soggetti all'azione delle cause costanti e delle variabili regolari, ma risentono altresì l'influenza delle cause variabili irregolari, perturbatrici o accidentali, che li fanno deviare in ogni senso in più o in meno. Bisogna dunque scoverarli da tutto ciò, che possono avere di accidentale e di anormale, sicchè altro non rappresentino che il prodotto delle cause costanti o delle variabili regolari „.

Alle *medie* si sono attribuite molte qualità e se ne sono date molte definizioni. Preferibile quella di Ferraris: Il va-

lore medio o la *media* è una quantità derivata dalle altre quantità della serie compresa fra la massima e la minima di essa, e ottenuta in modo che tutte le quantità della serie la determinano in ragione della loro grandezza e del loro numero „.

La *media* è l'espressione numerica che riassume le condizioni dei fenomeni omogenei non tipici.

La *media*, si afferma, sintetizza e compendia in sé tutta una serie di termini; essa contiene ciò che di costante e di comune esiste negli elementi di una serie. Si dice che nella *media* scompare ciò ch'è variabile ed accidentale, per mettere in evidenza ciò che è il prodotto della *cause costanti*; che la *media* riduce gli errori e le deviazioni alle minime proporzioni e rappresenta il valore più probabile, più prossimo al vero.

36. *Varie specie di medie.* — La prima distinzione da farsi è tra *media tipica* e *media indice*.

Si dice *tipica*, secondo Quetelet, il *valore* o la *grandezza* intermedia di una serie d'individui tendenti ad un solo tipo.

In conformità della precedente osservazione del Salvioni la *media tipica* è il valore dal quale poco si discostano in più o in meno i singoli termini dai quali è tratto e attorno al quale si dispongono simmetricamente e con frequenza tanto minore quanto più si allontanano dalla *media* stessa.

Così distribuendo 10,000 Italiani secondo l'indice cefalico, si avrebbe intorno alla *media* 82,7 una ripartizione simmetrica, una frequenza di casi tanto più decrescente, quanto più ci discostiamo dalla *media* per andare verso gli indici cefalici più bassi o verso i più elevati, come risulta da questi dati:

<i>Indice cefalico :</i>	<i>Individui osservati :</i>
70	12
da 70 a 72	103
„ 73 „ 75	463
„ 75 „ 78	1,261
„ 79 „ 81	2,020
„ 82 „ 84	2,702
„ 85 „ 87	1,898
„ 88 „ 90	1,086
„ 91 „ 93	347
„ 94 „ 96	96
Oltre 96	12
Indice medio 82.7	10,000 (<i>Benini</i>).

Se i dati da cui si ottiene la media sono di natura antropologica, essendo essa tipica, se ne può indurre che gli individui osservati appartengono alla stessa razza o ad una collettività con minima miscela di razze; se si tratta di una media di fenomeni sociali, se essa sarà tipica, si può supporre, che la collettività abbia una composizione molto omogenea per le condizioni sociali dei suoi membri.

La media si dice *indice* o *non tipica*, quando i dati dai quali si trae non presentano la disposizione sopraindicata e sono molti distanti tra loro, senza alcuna tendenza ad addensarsi maggiormente attorno ad un gruppo, come nel caso degli indici cefalici sopra esposto, nel quale sopra 10,000 individui 2,702 la presentavano da 82 a 84. Così, per la media della statura in Italia, se si prende la media dei soli soldati della provincia di Udine, dove si trovano le più alte stature colla proporzione massima del 29,63 % di statura sopra m. 1,70 — o quella dei soli soldati della provincia di Cagliari, dove si ha la minima proporzione di alta statura di oltre m. 1,70 del 5,63 %, si avranno due *medie tipiche* di statura. Ma la media della statura di tutti gli Italiani, che risulterà dalle più alte del settentrione, dalle meno alte del centro e dalle più basse del mezzogiorno, della Sicilia e della Sardegna, sarà una *media indice* e *non tipica*. Parimenti, se in un paese troveremo moltissimi piccoli proprietari e pochissimi proletari e latifondisti, la media proprietà di 20 ettari per abitante sarà *media tipica*. Se la stessa media di 20 ettari risultasse dalla esistenza di moltissimi proletari, di pochissimi piccoli proprietari e di un discreto numero di latifondisti, noi avremmo una *media indice*. Le medie dei più importanti fenomeni sociali — ricchezza, consumi, ecc. — sono quasi sempre *medie indici*, specialmente nelle grandi città; si avvicinano di più alle *medie tipiche* tra le popolazioni rurali.

Il punto di maggiore addensamento, che Bowley chiama *mode*, il Benini lo chiama *norma*. La seriazione sarà *uninormale*, se presenta un solo punto di massimo addensamento; *plurinormale*, se ne presenta due o più. La distribuzione della statura e degli indici cefalici in Italia ci presenta due addensamenti verso la più alta e la più bassa statura, verso la brachicefalia e la dolicocefalia; d'onde l'induzione, che per lo meno siano due le razze che la popolano.

La media poi si dice *obbiettiva* e *subbiettiva*, *reale* e *ideale*. La *media obbiettiva* corrisponde alla *reale* ed al *mean* degli Inglesi; la *subbiettiva* all'*ideale* o all'*average* degli stessi Inglesi.

La *media obbiettiva o reale o mean*—detta anche *probabile* dal Messedaglia—corrisponde alla media aritmetica di diverse misure di uno stesso oggetto. Esempio: si misura più volte la statura di un uomo, di un campanile, ecc., e si trova che ogni misura differisce dall'altra ora in più, ora in meno. Le differenze derivano o dalle condizioni dell'osservatore o dalle imperfezioni dello strumento, che si adopera ecc. La *media* di queste misure diverse sarà *obbiettiva, reale, probabile (mean)*.

La *media subbiettiva ideale o average* è quella, che risulta dalle misure di diversi oggetti omogenei, della stessa qualità. Esempio: prendo la misura della statura di 1000 Italiani della stessa età appartenenti alle diverse regioni; la statura media, che ne risulterà, sarà la *media subbiettiva o ideale o average*.

Giustamente, però, osserva il Salvioni, che la denominazione di *sogettiva* non è la migliore, perchè *sogettivo* vuol dire di ordinario che si riferisce all'uomo, allo statistico che osserva. E' preferibile la denominazione di *media ipotetica* appoggiandosi essa all'ipotesi delle cause costanti, di cui sarebbe l'espressione.

La *media subbiettiva* può essere *tipica o indice*.

In generale si dice che una *media*, perchè possa rappresentare con approssimazione il valore che si cerca, deve soddisfare principalmente ad alcune condizioni.

Dev'essere dedotta; a) da termini possibilmente omogenei per unità di oggetto, di tempo, di luogo; b) da termini individuali, piuttosto che da termini medii; poichè la media delle medie può non coincidere colla media dei termini singoli, da cui le medie sono state dedotte; c) da un numero grande di osservazioni, in guisa che le variazioni in un senso possano compensarsi con quelle in senso contrario; d) da singole serie parziali, se la serie intera è di carattere dinamico, poichè la media della serie generale non può farci scoprire la tendenza del movimento e lo stesso movimento (*Gabaglio*).

37. Dal punto di vista matematico si distinguono quattro specie di medie: *aritmetica, geometrica, armonica ed antiarmonica*.

Media aritmetica: viene data dalla *somma dei termini di una serie divisa pel numero dei termini stessi*.

Si abbia una serie di dati rappresentati dai numeri 2, 5, 9, 12: la loro media è data dalla somma divisa per quattro

$$\frac{2 + 5 + 9 + 12}{4} = 7.$$

E' bene osservare, per quanto si dirà in appresso, che la stessa *media* 7 può ottenersi da un numero maggiore o minore di termini, come in questo caso:

$$\frac{1+4+10+7+2+8+3+19+5+11}{10} = 7.$$

Media geometrica: viene data dalla *radice del prodotto dei termini*. Se i termini saranno due, ad esempio 2 e 6, si avrà la *media*.

$$\sqrt{2 \times 6} = 3,494.$$

Media armonica: si ottiene *colla divisione del proprio prodotto dei termini osservati per la somma dei termini stessi*.

$$\text{La media armonica di 2 e 6} = \frac{2(2 \times 6)}{2 + 6} = 3.$$

Media antiarmonica: viene data *dalla somma dei quadrati dei termini divisa per la somma dei termini stessi*.

$$\text{La media antiarmonica di 2 e di 6} \dot{=} \frac{2^2 + 6^2}{2 + 6} = 5.$$

La forma algebrica delle quattro medie è la seguente:

$$\text{Media aritmetica. } M = \frac{a+b}{2}$$

$$\text{Media geometrica. } M = \sqrt{a \cdot b}$$

$$\text{Media armonica. } M = \frac{2ab}{a+b}$$

$$\text{Media antiarmonica. } M = \frac{a^2 + b^2}{a+b}$$

38. *Stato normale quantitativo. Errore o scostamento dalla media*. — La media aritmetica è quella adoperata più comunemente e serve a determinare lo *stato normale quantitativo* del fenomeno. Lo *stato normale quantitativo* sarebbe quello in cui un fenomeno, essendo determinato soltanto da *cause*

costanti, non presenterebbe variazioni quantitative tra un anno e l'altro. Ma siccome c'è sempre l'intervento di altre cause variabili e accidentali, così anche la manifestazione da un anno all'altro presenterà delle variazioni quantitative. Quando queste variazioni sono assai piccole, allora la media si dice che rappresenta lo *stato normale quantitativo*, non ostante tali variazioni. Si abbia, ad esempio, la cifra rappresentante la somma delle *medie* annue di matrimoni per 1000 abitanti in Italia, dal 1872 al 1897, e sia di 20,043; divisa pel numero degli anni di osservazione (26), darebbe la *media* annua di 7,7; questa rappresenterebbe lo *stato normale quantitativo* dei matrimoni in Italia, perchè la media differirebbe poco dai termini, cioè dal numero annuo effettivo dei matrimoni in tutti i 26 anni di osservazione. Infatti noi avremmo una massima differenza *in meno* nel 1880, quando discesero a 6,97 ed una massima differenza *in più* nel 1875, quando salirono a 8,42.

Non si può parlare di *stato normale quantitativo* che nei fenomeni, la cui osservazione per un dato periodo di anni forma una *serie statica*.

La differenza *in più* o *in meno*, che si osserva tra la *media* o i singoli termini, che contribuiscono a formarla, si chiama *errore* o *scostamento*.

La parola *errore*, se si tratta di *media subbiettiva*, non indica mai un errore materiale di rilevazione nel senso ordinario della parola, nel senso di sbaglio. L'*errore* nelle medie, che si chiama *errore di osservazione*, in opposizione all'*errore di rilevazione* o sbaglio *materiale* rappresenta: la *deviazione*, lo *scostamento*, che la manifestazione quantitativa del fenomeno presenta per effetto delle cause variabili ed accidentali, dallo *stato normale quantitativo*, che sarebbe stato prodotto, se avessero agito soltanto le cause costanti.

L'*errore* o *scostamento* si chiama *probabile*, quando la sua grandezza non fu sorpassata in una metà delle operazioni e lo fu invece nell'altra metà.

Si avverta, però, che l'*errore* nella *media obbiettiva* o *reale* rappresenta un vero errore o sbaglio nelle rilevazioni. Si fa la *media*, infatti, di ripetute misure di un oggetto, perchè si è osservato che ogni misura differisce dall'altra in più o in meno, attribuendosi il fatto o alla inesattezza nell'osservazione o alla imperfezione dello strumento adoperato.

In questo caso soltanto si può dire che la *media* rappresenta il valore più *probabile*, ossia più vicino al *vero*, fra quelli che risultano da osservazioni ripetute, secondo la definizione del Gabaglio. Le frasi: *più probabile* e *più vero* sono inesatte,

riferite al valore della *media subbiettiva*. Infatti: il numero minimo di matrimoni di 6,97 nel 1880 non è forse altrettanto *vero e certo* quanto quello di 8,42 nel 1875? Ciò che c'è di non *vero* invece, nel senso ordinario delle parole, è la *media* di 7,7, poichè in tutti i 26 anni dell'osservazione non una sola volta il numero dei matrimoni corrispose realmente a tale cifra.

39. *Media aritmetica ponderata o composta*. — Parlando delle *serie*, si è vista la distinzione, che c'è da fare tra *serie a peso costante* e *serie a peso variabile*. La maggior parte dei fenomeni, che si studiano e di cui si ricercano le *medie*, sono a *peso variabile*. Ora la *media aritmetica semplice*, come si denomina quella sin ora presa in esame, non può darci un concetto adeguato del valore intermedio; bisogna ricorrere alla *media aritmetica composta* o *ponderata*, che deve tenere conto del *peso*.

Questo bisogno si comprenderà meglio con un esempio pratico. Si voglia conoscere il prezzo medio delle varie qualità di frumento comprato o del frumento comprato in diversi giorni della settimana. Se la *quantità* di frumento, che rappresenta il numero delle osservazioni o il *peso*, è uguale per ogni qualità o per ogni giorno, si può adoperare la *media aritmetica semplice*.

Ettol.	50	qual.	A	o nel giorno di	Lunedì	al prezzo di	L.	25
"	50	"	B	"	Mercoledì	"	"	22
"	50	"	C	"	Giovedì	"	"	24
"	50	"	D	"	Sabato	"	"	27

In questo caso, nel fenomeno *compra di frumento* varia un solo aspetto, quello del prezzo: il *peso* del fenomeno, cioè la quantità acquistata nei diversi giorni o delle diverse qualità è *costante*; perciò per conoscere il prezzo medio, basta sommare i quattro prezzi diversi e dividerli pel numero dei termini e si avrà il prezzo medio, che sarà di L. 24,50 per ettolitro.

Ma se, invece, le quantità comprate fossero diverse ed a prezzo diverso in questa misura:

Ettol.	50	qual.	A	o nel giorno di	Lunedì	al prezzo di	L.	25
"	25	"	B	"	Mercoledì	"	"	22
"	35	"	C	"	Giovedì	"	"	24
"	90	"	D	"	Sabato	"	"	27

allora non basta più una *media semplice*, perchè il *peso* (la quantità comprata per ogni giorno o qualità) è diverso per

ogni prezzo. Si deve fare perciò la *media aritmetica composta* o *ponderata*, nella quale si tiene conto del *peso*. La si ottiene moltiplicando il prezzo di ogni compra per la sua quantità, sommando i singoli prodotti e dividendoli per la somma delle quantità comprate. Il prezzo *medio*, tenuto conto del *peso*, si avrà così:

$$\frac{50 \times 25 + 25 \times 22 + 35 \times 24 + 90 \times 27}{50 + 25 + 35 + 90} = \frac{5070}{200}$$

Il prezzo medio in questo caso sarà di L. 25,35. Si deve ricorrere alla *media ponderata* quando le condizioni *variabili* del fenomeno sono più di una. Nel caso sopracitato variava la *quantità* e il *prezzo* del grano acquistato.

Il Bowley adduce parecchi esempi importanti, per la determinazione dei salari medi secondo i vari modi di segnare il *peso*, mostrando le differenze, che ne risultano; e si comprende agevolmente la diversa importanza sociale ed economica, che ha la determinazione del salario *medio* dei lavoratori ottenuto con la *media aritmetica semplice* o colla *media aritmetica ponderata*, quando pochissimi lavoratori scelti hanno salari abbastanza elevati, mentre la grande massa ha dei salari bassissimi.

Sopra 1000 operai, ad esempio, ce ne siano 75 con un salario settimanale di L. 60; 125 di L. 40; 300 di L. 35 e 500 di L. 20. Colla *media aritmetica semplice*, il salario *medio* dei 1000 operai sarebbe dato da:

$$\frac{60 + 40 + 35 + 20}{4} = \text{L. } 38.75.$$

Coll'altra media, invece, si avrebbe:

$$\frac{75 \times 60 + 125 \times 40 + 300 \times 35 + 500 \times 20}{75 + 125 + 300 + 500} = \frac{30000}{1000}$$

Il salario *medio* in questo caso discenderebbe a L. 30.00.

Le medie più adoperate sono quella *aritmetica (semplice o ponderata)* e quella *geometrica*; prevale l'uso della prima, chiamata perciò *media statistica* per eccellenza.

La *media aritmetica* e la *geometrica* però non si adoperano indifferentemente, la scelta dell'una o dell'altra viene determinata dai rapporti di grandezza tra i valori, da cui la si vuole trarre. Nel caso più frequente, tali rapporti si svolgono senza legge o senza legge conosciuta o con una legge di cui non occorre tener conto. In tali casi si adopera la *media aritmetica*.

Bisogna ricorrere alla *media aritmetica*, secondo il Messedaglia, quando il caso mostra di procedere per somme o differenze; alla *geometrica*, quando, invece, è quistione di prodotti o quozienti, o, in generale, di rapporti per moltiplicazione o divisione.

Del criterio nella scelta della media lo stesso Messedaglia dà alcuni esempi. "Se si tratta della ragione media annuale d'incremento di un capitale, la media da' prendere è l'aritmetica: si suppone che il capitale aumenti ad *interesse semplice*, ossia per l'addizione successiva di una certa quantità costante; si ricorre alla *geometria*, se ad *interesse composto*, ossia quando l'incremento successivo si effettua in proporzione di quello già bello e avvenuto e del montante raggiunto. In generale, poi, la *media aritmetica* si deve adoperare nelle *serie a carattere statico*; quella *geometrica*, nelle *serie a carattere dinamico*.

40. *Massimi e minimi*. — La media aritmetica di una serie di termini, che si dà per molti fenomeni economici e sociali, può risultare da termini molto diversi tra loro per la grandezza e in guisa che alcuni termini presenteranno un *errore* o *scostamento* in più o in meno dalla media, molto sensibile. La media stessa può essere inoltre il prodotto di termini più o meno numerosi. Ad esempio: 10 può rappresentare la media aritmetica di

$$\frac{11 + 9}{2} \quad \text{e di} \quad \frac{24 + 5 + 1}{3}$$

In questi casi, il lettore potrà formarsi un concetto assai erroneo del fenomeno, come se lo formerebbe sulla distribuzione della ricchezza espressa dalla media di un paese, ottenuta colla somma della ricchezza di pochi milionari e di moltissimi nullatenenti o che posseggano una proprietà piccolissima; la media potrebbe riuscire discreta e farebbe supporre una condizione di generale agiatezza assai lontana dalla realtà.

In questi casi, si somministrano delle indicazioni, sul valore de' termini che concorrono a formare la media, che possono dare un concetto più adeguato ed approssimativo al vero del fenomeno esaminato.

Per riuscire nell'intento, si possono mettere accanto alla *media* i valori estremi: i *massimi* e i *minimi termini*, che contribuirono a produrla.

Si possono presentare accanto alla media di tutti i termini le *medie de' massimi* e *de' minimi*. La prima in questo caso sarà detta *media generale*.

41. *Media di gruppi di una serie.*—La *media generale* di una lunga serie di anni nasconde il decorso del fenomeno, e non lascia scorgere all'osservatore se esso presenta la tendenza ad aumentare o a diminuire, e non si può quindi stabilire se abbiamo da fare con una *serie statica* o *dinamica*, e nel secondo caso, se c'è tendenza allo aumento o alla diminuzione. In questo caso, la serie si divide in parecchi gruppi, o in serie, che comprendono un minor numero di quantità statistiche. La divisione, però, non dovrà essere arbitraria; ma si farà in guisa, che ogni gruppo comprenda un egual numero di anni; così, se la serie comprende le osservazioni di 30 anni: potrà dividersi in tre gruppi di 10 anni per uno o in 6 di cinque per uno, ecc. La divisione in gruppi potrà non essere fatta a periodi uniformi, purchè essi siano *naturali*, dice il Salvioni, cioè dati dalle fasi reali del fenomeno, che si considera, e dalle cause, che influiscono sopra di esse.

Però anche quando c'è un fenomeno, che ha una importanza eccezionale, la serie si può dividere in gruppi uniformi che hanno, per così dire, come centro quel dato fenomeno.

Si sa, ad esempio, che le tariffe doganali generali votate dal Parlamento italiano nel 1887 ed entrate in vigore il 1.º Gennaio 1888 esercitarono una forte perturbazione nel movimento commerciale del nostro paese: è noto del pari, che uno dei prodotti della nostra esportazione, che maggiormente se ne risentì fu il vino.

In questo caso, c'è un fenomeno, che deve servire di criterio nella divisione per periodi della intera serie del movimento commerciale; ad esempio, nel ventennio 1881-910, il mutamento di regime doganale avvenuto il 1.º gennaio 1888. Perciò, chi vuole statisticamente porre un paragone esatto, deve dividere in gruppi omogenei il periodo anteriore 1871-87 e quello posteriore 1888-910. La divisione, in questo caso non può assolutamente essere arbitraria.

Con queste divisioni, le *medie* de' singoli gruppi potranno dare la indicazione del decorso del fenomeno colla rispettiva tendenza all'aumento e alla diminuzione.

Il bisogno di dividere una serie in diversi gruppi risulta evidente dall'esame della mortalità italiana dal 1871 al 1900. La mortalità dell'intero periodo sarebbe in media di 27,2 per ogni anno, e sarebbe molto elevata. Dividendo l'intero periodo in cinque quinquenni, per ogni quinquennio si avrebbero le seguenti medie:

1871-75	per 1000 abitanti morti	30,5
1876-80	" " "	29,5
1881-85	" " "	27,3
1886-90	" " "	27,4
1891-95	" " "	25,4
1896-900	" " "	22,9

Questa divisione in gruppi della serie ci apprende: 1.º che l'intera serie è a *carattere dinamico*: 2.º che la serie presenta la *tendenza* alla diminuzione costante.

42. *Valore mediano, media di densità ed ordinata massima.* Accanto ai *valori medi*, che sono quelli rappresentati dalla media aritmetica o geometrica, si ricercano altri valori, che possano sostituire i primi o integrarli e far intendere meglio il decorso della serie e la grandezza dei termini, che la compongono.

Un valore, che si trova facilmente è il *mediano*, che viene dato dal termine, che sta nel mezzo di una serie, i cui elementi non sono disposti cronologicamente, ma in ordine decrescente, dal più grande al più piccolo: o viceversa.

Un esempio più chiaro del *valore mediano* lo si ricerchi nelle stature. Avendo 151 soldati di diversa statura si dispongono in ordine decrescente. La statura del 76.º soldato che si troverebbe nel mezzo, sarebbe la *statura mediana*.

Questa statura mediana, però, potrà essere molto distante dalle altre stature o in più o in meno, a seconda che i gruppi etnici, che comporranno la serie, saranno più o meno numerosi.

Si sa, ad esempio, che alcune provincie del Veneto presentano stature alte: medie, quella della Italia centrale; bassissime, quelle della Sardegna. Ora, se nella serie di 151 soldati ve ne saranno 19 delle provincie venete ad alta statura; 57 delle centrali con stature medie e 75 delle sarde con stature basse, è evidente che il 76.º, che servirà a dare il *valore mediano*, sarà molto distante dai 75 Sardi e dai 19 Veneti.

Nello studio dei fenomeni sociali, più che la *media statistica* o il *valore mediano*, hanno maggiore importanza la *media di densità* e l'*ordinata massima*.

Per ottenere la *media di densità*, i termini della serie si dispongono in ordine decrescente, nè più nè meno come per ottenere il *valore mediano*. Si vede attorno a quale cifra si aggruppa un maggior numero di termini; la media di questi termini è la *media di densità* ed indica che la manifestazione quantitativa del fenomeno più spesso vien data da questa media. La quale, a mio avviso, meglio che la *media statistica*

ordinaria, può rappresentare lo stato normale quantitativo del fenomeno.

L'aggruppamento dei termini deve avvenire con criteri fissi e determinati, separando i gruppi gli uni dagli altri seguendo una data misura o scala o *modulo*: di una mezza unità, di una unità, di cinque unità, ecc. ecc. tra un gruppo e l'altro.

Trattandosi di fenomeni, che non sono molto numerosi e le cui oscillazioni estreme non distano molto tra loro, come ad esempio, il numero dei matrimoni per 1000 abitanti, che in Italia oscillano pel 1872-900 attorno alla media di 7,27, con un massimo di 8,42 nel 1875 ed un minimo di 6,88, nel 1898, ogni gruppo si differenzierà dall'altro di mezza unità.

Come si ottenga tale *media di densità* risulterà all'evidenza guardando alle cifre dei matrimoni in Italia dal 1872 al 1900 disposte, non cronologicamente, ma in ordine decrescente.

MATRIMONI PER 1000 ABITANTI

Anni:	da 6,50 a 6,99	Anni:	da 7,50 a 7,99
1898 . . .	6,88	1872 . . .	7,53
1880 . . .	6,97	1879 . . .	7,60
	da 7,00 a 7,49	1874 . . .	7,64
1896 . . .	7,07	1889 . . .	7,67
1878 . . .	7,17	1877 . . .	7,76
1900 . . .	7,19	1882 . . .	7,84
1897 . . .	7,22	1886 . . .	7,93
1895 . . .	7,29	1873 . . .	7,94
1899 . . .	7,33	1888 . . .	7,94
1890 . . .	7,34	1887 . . .	7,96
1893 . . .	7,39		da 8,00 a 8,50
1892 . . .	7,45	1885 . . .	8,01
1894 . . .	7,45	1883 . . .	8,06
1891 . . .	7,48	1881 . . .	8,11
		1876 . . .	8,18
		1884 . . .	8,26
		1875 . . .	8,42

Che cosa ci dice questa disposizione? Che i matrimoni per due anni oscillarono tra 6,50 e 6,99, per 1000 abitanti; per undici anni tra 7,00 e 7,49, per dieci tra 7,50 e 7,99; per sei tra 8,00 e 8,49. La *media di densità* verrà data quindi dalla *media aritmetica* del secondo gruppo, ch'è più numeroso di tutti gli altri tre, e che rappresenta le oscillazioni tra 7,00 e 7,48. Questa media sarà di 7,30.

In questa guisa noi avremmo questi tre valori intermediari per i matrimoni in Italia dal 1872 al 1900:

Media aritmetica.	7,27
Valore mediano (anno 1879).	7,60
Media di densità	7,30 *

La *media di densità*, come si vede, è più alta della *media aritmetica* e più bassa del valore *mediano*, ed essendo tratta dalla manifestazione quantitativa del fenomeno per un maggior numero di anni — 11 su 29 —, più probabilmente rappresenterà lo *stato normale quantitativo del fenomeno*.

Ed ora più brevemente della *ordinata massima*, che rappresenta quel dato, che comprende il maggior numero di termini di ugual valore, Si abbiano 1000 soldati di diversa statura cogli estremi di oscillazioni tra 180 e 145 centimetri, e siano così distribuiti:

3	di metri	1,80
5	”	1,75
10	”	1,70
15	”	1,65
25	”	1,62
12	”	1,60
15	”	1,55
10	”	1,50
5	”	1,45
<hr/>		
100		

L'*ordinata massima* delle stature verrà data da 1 metro e 62 centimetri, perchè il gruppo più numeroso di soldati, 25, presentava quella altezza.

L'*ordinata massima* può servire meglio della *media ponderata* per indicare la frequenza e l'importanza di alcuni fenomeni sociali, ad esempio, quello dei salari: e la dimostrazione della importanza della divisione tra le varie classi di operai con salari diversi si avverte più chiaramente per mezzo di una rappresentazione grafica — di un diagramma.

43. *Metodo seriale*. — Tutti i vari mezzi escogitati sinora per attenuare, se non eliminare interamente, gl'inconvenienti delle *medie*, non raggiungono interamente lo scopo. La natura, l'intensità e la distribuzione del fenomeno che si studia.

* Le cifre riportate sono medie delle medie annuali.

però, si conosceranno meglio ricorrendo al cosiddetto *metodo seriale*. Fu proposto e adoperato pei primi da Morselli e Le Bon nell'antropologia; ma riesce anche utile e di larga applicazione nelle scienze sociali.

Col *metodo seriale* " le quantità ottenute colla rilevazione si distribuiscono su una scala distinta per gradi uguali dal *minimum* al *maximum*. Il grado della scala dei valori in una seriazione dicesi *modulo*, e poichè ne risultano tanti gruppi numerici, il modulo è anche detto di *aggruppamento*. Qualunque sia la serie numerica presa in esame, la distribuzione dei valori mediante il modulo avviene sempre per una scala ascendente da un lato o discendente dall'altro lato della media o almeno della mediana; ad ogni grado corrisponde un gruppo di variazioni individuali o di scarti (*errori*). Si ha colla seriazione il vantaggio di scorgere tutto il campo *possibile* degli errori e delle deviazioni e nello stesso tempo la probabilità di ciascuno di essi. Di fronte al quale vantaggio appare anche maggiore l'artificiosità fallace della semplice media aritmetica „ (*Morselli*).

Meglio che da questa descrizione, che cosa sia il *metodo seriale* e quale la sua superiorità su quello delle *medie* risulterà da qualche esempio. Si voglia conoscere, ad esempio, com'è distribuita in Francia, la ricchezza, desumendolo dalle successioni.

Nel 1905, le successioni dichiarate e nelle quali l'attivo superava il passivo furono 385,019 per l'ammontare di 5 miliardi, 745 milioni, 888,713 lire. Col sistema della *media aritmetica* si potrebbe stabilire che ciascuno dei 385,019 morti lasciò una eredità di poco più di L. 14.926. Ne risulterebbe una discreta condizione di tutti. Aggruppando le successioni col *metodo seriale*, si riesce a ben altro risultato.

Edgard Milhaud seguendo il metodo di De Foville in base all'ammontare delle suddette successioni e donazioni ha calcolato che la ricchezza privata della Francia sia di lire 207,999,000.000. Ripartita questa tra i 39 milioni (in cifra tonda) di abitanti della Francia col sistema della *media indice* sembrerebbe che ciascun Francese possedesse una ricchezza media di poco più di L. 5,333. Ma in base al numero di coloro, che morirono senza lasciar nulla, e degli altri, che lasciarono un patrimonio, calcolò che la ricchezza totale era divisa soltanto tra 12,293,956 individui. Ripartendo tra loro la ricchezza totale, si avrebbe che circa 27 milioni sarebbero nullatenenti, e i 12,295,035 avrebbero una ricchezza media di poco più di L. 16,918.

Ma nemmeno questa media darebbe un'idea della realtà:

alla quale ci avviciniamo aggruppando col *metodo seriale* gli individui secondo la loro ricchezza. Con questo aggruppamento il Milhaud ha trovato, sempre in base all'ammontare delle singole successioni, questa distribuzione:

Numero individui	Capitale posseduto da ogni gruppo.	Media per ciascun membro del gruppo.
61	5.241.600.000	85.927.868
276	5.574.400.000	20.193.156
768	6.344.000.000	8.260.416
3.688	14.185.000.000	3.846.250
10.327	18.761.600.000	1.816.752
22.991	20.072.000.000	912.000
49.302	92.131.200.000	444.428
142.621	28.454.400.000	199.540
227.333	20.198.800.000	90.431
1.337.812	37.440.600.000	27.978
4.295.413	26.671.600.000	6.566
3.337.585	6.385.600.000	1.913
3.867.759	1.538.200.000	3.098
12.295.936	207.909.000.000	

Con questo aggruppamento *seriale* risulta che 27,000,000 cioè il 69% circa della popolazione francese nulla possiede; e che 15,120 cioè 0,038% della popolazione possiedono 50 miliardi e 106 milioni, cioè il 24,08% della ricchezza totale. La sproporzione nella distribuzione della ricchezza non potrebbe essere più spaventevole.

Valore delle successioni in Italia per l'anno fiscale 1904-905.

Ammontare della successione		Numero delle successioni	
Da L.	1 a L.	500	51.171
" "	500 " "	1000	25.085
" "	1000 " "	2000	21.507
" "	2000 " "	4000	16.549
" "	4000 " "	10.000	13.612
" "	10.000 " "	50.000	9.080
" "	50.000 " "	100.000	1.388
" "	100.000 " "	300.000	590
" "	300.000 " "	300.000	150
" "	500.000 " "	500.000	110
Oltre L.	1.000.000		51

Qui l'aggruppamento è fatto secondo una data scala, che ci deve essere sempre quando si ricorre al *metodo seriale*.

I gruppi, però, non sono divisi con una scala uguale per ragione di comodità. Il quadro sarebbe divenuto lunghissimo, se si fosse ammesso per modulo di aggruppamento L. 500; ciò facendo, i gruppi non dovevano essere da L. 1 a 500, da L. 501 a 1000, da 1001 a 2000, ecc., ma da 1 a 500; da 501 a 1000; da 1001 a 1500; da 1501 a 2000, ecc., ecc.

È evidente che nel *metodo seriale* non si esclude l'uso dei valori intermediari e degli aggruppamenti; per escluderli e rappresentare il fenomeno in tutta la sua precisione, si dovrebbero esporre i singoli casi e aggruppare soltanto quelli, che rappresentano quantità perfettamente uguali.

Col *metodo seriale*, nello studio dei fenomeni sociali, si coglie più facilmente l'importanza ed il significato della *media di densità* o della *ordinata massima* del fenomeno. Di questo *metodo seriale* si sono fatte varie applicazioni dal Bowley, dal Mayo-Smith, dal Booth e da altri.

44. *Critica delle medie e dei valori mediani.* — Le medie e i valori mediani sono stati fatti segno sempre alle più aspre critiche. Si osservò, che essi non rappresentano la realtà: rappresentano invece un letto di Procuste, sono illusioni ottiche, prodotti fantasmagorici, ecc.

Contro le *medie* si obietta, che esse non corrispondono alle *idee generali*, perchè le idee si riferiscono alle *qualità* e la *media* alle *quantità*; che le *idee generali* risultano dalla diminuzione dei caratteri differenziali, mentre le *medie* si ottengono sottraendo da un lato e aggiungendo dall'altro; che le *idee generali* abbracciano tutti i casi particolari, mentre la *media* varia per ogni caso di addizione o di sottrazione; l'addizione o la sottrazione di un solo termine, anzi, può spostare sensibilmente la *media*, se questo solo termine rappresenta una quantità importante; che l'idea corrisponde a qualche cosa di comune, d'invariabile — ad esempio: l'*idea* di cavallo — e la *media* a quantità diverse, mutabili — ad esempio: l'*età media*, che può esprimere condizioni demografiche assai diverse.

Infatti, ad esempio: l'età *media* di 30 anni di una popolazione può essere il prodotto di un numero esuberante di fanciulli e di vecchi deboli, che consumano e non producono; può risultare da una prevalenza dell'età degli adulti in età *media*, cioè produttori e vigorosi.

La *media*; se risulterà da numerose osservazioni e su vasta superficie, renderà molto difficile la ricerca delle cause.

Prendiamo, dice il Pareto, la durata *media* della vita: se essa aumentasse in alcune provincie e diminuisse in alcune altre, prendendo la *media* per l'intero Stato, si troverebbe

stazionaria; ovvero mutata in più o in meno; ma sempre in proporzioni diverse dalla realtà. Come si studierebbero le cause locali dell'aumento o della diminuzione della mortalità?

Non va dimenticato, infine, che tra gli stessi apologisti più eminenti delle *medie* talora non c'è accordo sulla *media*, che dovrebbe rappresentare il valore più probabile del fenomeno; alcuni preferiscono la *media aritmetica*; altri la *geometrica*, l'*armonica*, ecc.

In ogni modo, i servizi, che rendono le *medie*, sono innegabili e non piccoli. Senza le *medie*, noi non potremmo formulare ed emettere un giudizio su certi avvenimenti e sul decorso dei fenomeni sociali.

Tutto ciò che si riferisce alle cose umane non è perfetto e non è assoluto e dobbiamo spesso contentarci del possibile. Ora, a parte l'importanza che le *medie* hanno pel finanziere, pel commerciante, per le Società di assicurazioni, ecc., nella vita comune ci riuscirebbe difficilissimo, se non impossibile, afferrare la realtà stessa senza il sussidio dell'astrazione delle *medie*. Anche sfuggendoci i dettagli, riuscirà sempre utilissimo il sapere se, in *media*, in un dato paese aumenta o diminuisce la ricchezza, l'emigrazione, l'alcoolismo, la delinquenza, l'analfabetismo, ecc.

Questa conoscenza *media* ci permette in blocco di giudicare se una società progredisce o regredisce; quando avremo l'idea generale del fatto, potremo scendere ai dettagli suoi più precisi. Lo statistico, del resto, ha messo lo studioso in sull'avviso sul valore da assegnare alle *medie*; e gli ha anche indicati i modi, nella impossibilità dello studio individualizzato dei fenomeni atipici delle grandi collettività, di correggerne e di attenuarne gl'inconvenienti. L'uso sempre più largo e più perfezionato del *metodo seriale* renderà grandi servizi nella scienza e nella vita pratica*.

45. *Legge di distribuzione degli errori accidentali.* — Discorrendo delle varie specie di *medie* si disse, che la *media oggettiva* o *reale* è quella presa tra diverse, ripetute, osservazioni o misure di uno stesso oggetto, e che si ricorreva alla *media*, perchè tra le varie osservazioni o misure c'erano differenze quantitative o *errori* (parag. 37). Questa specie di errore, si osservò, nel paragrafo successivo (38), che non do-

* A chi volesse condannare senza appello le *medie*, di cui ho fatto più volte la critica — specialmente in: *Alcoolismo; sue conseguenze morali e sue cause*, lo statistico risponderebbe: indicatemi qualche procedimento migliore da sostituire.

veva confondersi coll' *errore statistico* o *scostamento* dalla *media*.

Intanto l'una e l'altra specie di errori si presentano con una certa regolarità, ch'è stata chiamata *legge di distribuzione degli errori accidentali*, colla quale si giustifica l'uso della *media*, che rappresenterebbe il valore più probabile dei dati dell'osservazione nella *media oggettiva* e lo *stato normale quantitativo* di molti termini, che concorrono a formare una *media subbiettiva* o *ideale*.

Si chiama *legge di distribuzione degli errori accidentali*, perchè si riferisce a quegli *errori statistici* o *scostamenti* prodotti dalle cause accidentali, che producono il maggiore o minore scostamento dallo *stato normale quantitativo* (ipotetico), che dovrebbe essere il risultato dell'azione delle *cause costanti*.

Col calcolo matematico, si determina come si distribuiscono tali errori in ordine alla frequenza e alla grandezza e si trova: 1.º che aumentando, il valore assoluto, la grandezza dell'errore, diminuisce il numero delle volte in cui l'errore stesso si manifesta e si riproduce; 2.º che sempre in gran numero di osservazioni, l'errore positivo e l'errore negativo si presentano con la stessa frequenza, lo stesso numero di volte. Così, ad esempio, dice il Benini, misurando il torace di un uomo, accadrà più facilmente di sbagliare di un solo centimetro (in più o in meno), che non di due, di tre o più. Se ne conchiude che la *frequenza relativa di un errore dipende dalla grandezza di esso*; ciò che, espresso in forma matematica equivale a quest'altra proposizione: *La probabilità dell'errore è una funzione dell'errore stesso; e vi è la medesima probabilità di commettere un errore positivo come un errore negativo dello stesso valore*.

“Prendendo come unità di misura lo *scarto* o *errore probabile*, che sappiamo quale sia (paragr. 38), la frequenza di qualsivoglia altro errore, più grande o più piccolo, resta teoricamente determinata ed è facile calcolarla coll'aiuto di apposite tavole di valori, che si trovano nei trattati di calcolo delle probabilità „.

Questa distribuzione degli errori accidentali, determinata da Quetelet, geometricamente assume forma campanulare — a cappello di carabiniere, come dice Tammeo — e fu chiamata dallo stesso Quetelet *curva binomiale*, perchè il suo andamento geometrico corrisponde allo sviluppo algebrico del binomio di Newton. Ciò che si rileva dalla seguente figura: