Rendiconti
Accademia Nazionale delle Scienze detta dei XL
Memorie di Matematica
102° (1984), Vol. VIII, fasc. 2, pag. 27-48

GIUSEPPINA VOLPATO MARTINELLI (*)

A proposito della caratterizzazione, discretizzazione e simulazione del processo poissoniano (**)

On the Characterization, Discretization and Simulation of the Poisson Process

SUMMARY. — A discrete-time stochastic process of arrivals of customers at a counter (in a figurative sense), in which the arrivals are related to a Bernoulli sampling scheme with two alternatives and which has recently been described by M. Volpato, is being modified here by considering the arrivals as related to a bloc-sampling scheme, always with two alternatives, thus freeing the arrivals from the assumption of independence. The new marginal laws that describe the waiting time for the arrival of the r-th customer, the time length between two successive arrivals (which also in this case characterizes the discrete process), and the total number of customers that may arrive in a fixed period of time, still converge (for the step length of the equipartition of the R^+ becoming infinitesimal) to the analogous marginal laws that are typical for the continuous-time Poisson process. By means of an appropriate flow-chart we describe an algorithm for the simulation of the marginal laws of the discrete process so as to make possible its application to those real situations, in which it may be useful not to loose any property (such as dependence or independence), that is being lost with the passage to the limit toward the continuous-time Poisson process.

In un recente lavoro (1) è stato definito un processo stocastico (di arrivi di clienti ad uno sportello cui viene richiesto un servizio) a tempi discreti, chiamato processo bernoulliano, ed indicato col simbolo $\mathbf{B}(\lambda/n)$, perché caratterizzato dall'ipotesi che in ciascuno dei tratti parziali di ampiezza $1/n$ $(n=1, 2, \ldots)$, di una equipartizione D_n del semiasse positivo dei tempi, possa arrivare *uno ed uno solo cliente* con la probabilità λ/n (pari a quella di ottenere il successo in una qualsivoglia estrazione di un processo di estrazioni bernoulliane da un universo a due alternative: successo ed insuccesso) il numero reale λ/n assumendo

(*) Indirizzo dell'A.: Istituto di Statistica dell'Università degli Studi, Cà Foscari di Venezia.
(**) Memoria presentata il 9 agosto 1983 da Giuseppe Scorza Dragoni, uno dei XL.
(1) M. VOLPATO, *Sulla caratterizzazione, discretizzazione e simulazione del processo poissoniano*, Rend. Accad. XL, Vol. VII (11), 177-192 (1983).

ISSN-0392-4106

il significato di numero medio di clienti arrivabili nell'unità di tempo, convenzionalmente assunta uguale ad $1/n$.

In questa nota dimostro che le conclusioni del lavoro cit. in (¹) continuano a sussistere anche quando, dall'universo a due alternative, si eseguono le estrazioni *a blocco* (cioè, col metodo *esaustivo* che prescrive il non reimbussolamento dell'estratto). Le variabili discrete ξ_n (pascalliana) ed η_n (bernoulliana) che in (¹) descrivono rispettivamente, il tempo d'attesa dell'r-mo cliente ed il numero complessivo di clienti arrivabili in un fissato periodo di tempo, a causa delle nuove ipotesi sul modo di effettuare le estrazioni, vedono ridursi l'insieme dei valori accessibili, e alterarsi le probabilità che assegnano ai medesimi nel senso che tali probabilità risultano ora moltiplicate per un fattore dipendente da n e convergente ad 1 al divergere di n. Nella nuova espressione la ξ_n è stata denominata una *pascalliana bloccata* (o *condizionata*) e la η_n (che sostanzialmente si identifica con una *ipergeometrica*) una *bernoulliana bloccata* (o, *condizionata*). Analoga alterazione subisce la variabile τ_n che descrive il tempo separatore di due arrivi successivi (peraltro, identificantesi con la ξ_n, relativamente al caso $r = 1$), che è stata chiamata *geometria bloccata* (o, *condizionata*) e che potrebbe rivelarsi di qualche interesse nei processi discreti di vita e di morte nei quali notoriamente si ricorre alla geometrica tradizionale.

Come dianzi si è accennato, facendo divergere n si ottengono anche stavolta le analoghe leggi marginali tipiche del processo stocastico a tempo continuo di Poisson. Tale processo, almeno per quanto attiene le leggi in discorso, può allora considerarsi *anche* come processo limite di un processo a tempi discreti, qui denominato, *ipergeometrico*. Per quanto riguarda le applicazioni vengono indicati anche questa volta diagrammi di flusso a grandi blocchi per la simulazione delle variabili marginali discrete che, appunto, caratterizzano il processo a tempi discreti nel quale gli arrivi dei clienti non sono indipendenti.

1. - Ipotesi caratterizzanti il processo di arrivi. Osservazioni

Eccole:

$\mathfrak{I}_1$) Si fissa una equipartizione D_n del semiasse positivo dei tempi in tratti parziali disgiunti

$$T_{i,n} \triangleq \left(t \in \mathbf{R}^+;\ \frac{i-1}{n} < t < \frac{i}{n}\right), \qquad (i = 1, 2, \ldots;\ n = 1, 2, \ldots)$$

di uguale ampiezza $1/n$ che viene assunta come unità di tempo;

$\mathfrak{I}_2$) Si predispone un'urna U_n, da considerarsi associata all'equipartizione D_n, contenente inizialmente N_n soggetti (da ritenersi equicomparibili in ciascuna estrazione anche nei vari stadi di un processo esaustivo di estrazioni) ripartiti in due classi disgiunte ed esaustive dette:

1) classe dei *successi* formata da $N_n p_n$ soggetti, con $p_n \doteq \lambda/n$, ove λ è un numero reale, positivo, eventualmente da approssimare col numero razionale $(n/N_n)k_n$, ove k_n è l'intero per il quale $k_n/N_n < \lambda/n < (k_n+1)/N_n$, in modo che $N_n p_n$ risulti in ogni caso un numero intero;

2) classe degli *insuccessi*, formata da $N_n q_n$ soggetti, con $q_n = (1-\lambda/n)$; il numero totale N_n dei soggetti soddisfacendo la relazione

$$\lim_{n\to+\infty} \frac{N_n}{n^2} \neq 0 \,,$$

determinante la divergenza del prodotto

$$N_n p_n = N_n \frac{\lambda}{n} = k_n \,,$$

indicante il numero dei soggetti qualificati come successi (*).

λ_3) Si ritiene che nel generico tratto parziale $T_{i,n}$ arriva *uno ed uno solo* cliente soltanto se *estraendo a blocco* da U_n un campione di taglia i $(i < N_n)$, si verifica che alla i-ma estrazione (cioè, nell'ultima estrazione da eseguire per completare il campione che, appunto, è di taglia i) compare un successo, oltre agli altri eventuali che possono essere comparsi nelle precedenti $(i-1)$ estrazioni.

Osservazione 1: Nell'i-ma estrazione (appunto, l'ultima occorrente per completare il campione) *ogni* soggetto rimasto nell'urna U_n ha una probabilità $1/(N_n - i + 1)$ di essere estratto e quindi di entrare nel campione. Tuttavia se non si conosce quanti sono stati i successi (e quindi, gli insuccessi) comparsi nelle precedenti $(i-1)$ estrazioni, non si può calcolare la probabilità del successo alla i-ma estrazione nel processo di estrazioni *a blocco*.

Tanto per esemplificare, ed anche perché utile nel seguito, se nelle precedenti $(i-1)$ estrazioni si sono avuti complessivamente $(r-1)$ successi, e

(*) Se N_n fosse un infinito di ordine uno (e non di ordine >2, come indica la relazione di limite), k_n convergerebbe. Infatti, la prima delle disuguaglianze

$$\frac{k_n}{N_n} < \frac{\lambda}{n} < \frac{k_n+1}{N_n}$$

implicanti la relazione

$$0 < \frac{n}{N_n} k_n < \lambda$$

insieme con la seguente altra

$$\lim_{n\to+\infty} \frac{n}{N_n} \neq 0 \quad \text{(finito)}$$

esprimente che N_n è infinito di ordine uno, porgerebbe appunto la convergenza di k_n. È opportuno invece che k_n diverga per consentire la possibilità di arrivo di clienti anche quando l'orizzonte temporale diverge nel futuro.

quindi $(i-1)-(r-1)=i-r$ insuccessi, la composizione dell'urna U_n dopo la $(i-1)$-ma estrazione, è tale da contenere, complessivamente, $N_n-(i-1)$ soggetti, dei quali

$N_n p_n-(r-1)$, appartenenti alla classe dei *successi*;

$N_n q_n-(i-r)$, appartenenti alla classe degli *insuccessi*.

Ne segue che, attesa l'equicomparibilità dei soggetti presenti nell'urna in occasione di ogni singola estrazione, la probabilità del successo alla i-ma estrazione è

$$\frac{N_n p_n-r+1}{N_n-i+1}. \tag{1}$$

Osservazione 2: Come si è già segnalato, la probabilità di successo, in occasione di ogni singola determinata estrazione di un processo esaustivo di estrazioni, è una *probabilità condizionata*: dipendente cioè dai risultati delle precedenti estrazioni. Tuttavia, prescindendo dai risultati delle precedenti estrazioni, si può calcolare anche una *probabilità a priori* (o, probabilità *indipendente*) che può avere la comparsa del successo in una generica qualsivoglia estrazione del processo *esaustivo* di estrazioni. Il criterio di probabilizzazione per un tale evento può essere quello del rapporto di Laplace. Si può osservare infatti che nella formazione di un campione di taglia generica r, si possono ottenere, complessivamente, $D_{N_n,r}$ campioni *ordinati* da considerarsi la totalità dei *casi* ugualmente possibili agli effetti della conta dei *casi* che alla ν-ma estrazione (del processo esaustivo di estrazioni) compaia il successo. I *casi* favorevoli (alla comparsa del successo alla ν-ma estrazione) sono dati da quei campioni *ordinati* che al ν-mo posto contengono il successo (e quindi che, nei precedenti $(\nu-1)$ posti contengono una qualsivoglia disposizione semplice, appunto di classe $(\nu-1)$ degli altri N_n-1 soggetti). Quindi se in ciascuna disposizione semplice delle $D_{N_n-1,\nu-1}$ aggiungiamo un posto e lo occupiamo ponendovi, in tutti i modi possibili, uno degli $N_n p_n$ successi, si ottengono tutti i casi favorevoli. Perciò il rapporto

$$\frac{N_n p_n D_{N_n-1,\nu-1}}{D_{N_n,\nu}}=\frac{N_n p_n(N_n-1)(N_n-2)\dots(N_n-\nu+1)}{N_n(N_n-1)(N_n-2)\dots(N_n-\nu+1)}=p_n \triangleq \frac{\lambda}{n} \tag{2}$$

evidenzia che $p_n=\lambda/n$ è la probabilità *a priori* (o, probabilità *indipendente*) di ottenere il successo in una qualsivoglia estrazione di un processo esaustivo di estrazioni da U_n. Ciò premesso poniamo fin d'ora la seguente

Definizione: Un processo di arrivi aleatori caratterizzato dalle ipotesi $\mathfrak{I}_1$), $\mathfrak{I}_2$), $\mathfrak{I}_3$) è detto *processo a tempi discreti ipergeometrico*, e rappresentato col simbolo $\mathfrak{I}_g(\lambda/n)$.

La motivazione della scelta del nome e del simbolo apparirà più chiara nel seguito.

2 - Studio della variabile discreta, che descrive il tempo d'attesa dell'r-mo cliente nel processo $J_p(\lambda/n)$

A norma dell'ipotesi λ_3) in ogni singolo tratto parziale $T_{i,n}$ non può arrivare più di un cliente. Quindi se denotiamo con ξ_n la variabile che indica il tratto aleatorio nel corso del quale si totalizza un numero r, $(1 < r \leq N_n p_n)$ (*), di arrivi, risulta evidente, che

$$P(\xi_n = T_{1,n}) = P(\xi_n = T_{2,n}) = \ldots = P(\xi_n = T_{r-1,n}) = 0 .$$

Per quanto attiene il calcolo delle probabilità

$$P(\xi_n = T_{i,n}) ; \qquad \text{con } i = r, r+1, \ldots$$

si può osservare che l'evento $(\xi_n = T_{i,n})$, con $i > r$ richiede il simultaneo verificarsi dei seguenti altri due eventi:

1) totalizzazione di $(r-1)$ arrivi nei primi $(i-1)$ tratti parziali (precedenti $T_{i,n}$), senza pretesa che la totalizzazione si verifichi in $T_{i-1,n}$: Tale evento è possibile con la probabilità

$$\frac{\binom{N_n p_n}{r-1}\binom{N_n q_n}{i-r}}{\binom{N_n}{i-1}} \tag{3}$$

pari a quella di ottenere $(r-1)$ successi in un campione di taglia $(i-1)$ estratto *a blocco* dall'urna U_n (associata all'equipartizione D_n) e nella quale, come si è detto, i soggetti qualificati come *successi* sono presenti inizialmente nella misura $N_n p_n$, mentre quelli qualificati come *insuccessi* sono $N_n q_n$, con $p_n = \lambda/n$ e $q_n = (1 - \lambda/n)$ e $\lim_{n \to +\infty} N_n/n^2 \neq 0$;

2) comparsa di un successo alla i-ma estrazione, sapendo che nelle precedenti $(i-1)$ estrazioni si sono già verificati $(r-1)$ successi. La probabilità di questo evento a norma della *Osservazione* 1, è

$$\frac{N_n p_n - r + 1}{N_n - i + 1} , \tag{4}$$

che appunto è la probabilità condizionata dall'evento precedente relativo agli $(r-1)$ successi nei primi $(i-1)$ tratti parziali.

(*) La limitazione superiore per r è più che altro apparente. Infatti, fissato per r un intero (anche comunque grande), attesa la divergenza del prodotto $N_n p_n$ (conseguente al fatto che N_n diverge con ordine almeno uguale a due e $p_n = \lambda/n$), si può sempre considerare una equipartizione D_n con n sufficientemente grande in modo che risulti $r < N_n p_n$.

Per il teorema delle probabilità composte, la probabilità con la quale nel corso del tratto parziale $T_{i,n}$ potrà totalizzarsi il numero r di arrivi è allora

$$(5)\qquad \frac{\binom{N_n p_n}{r-1}\binom{N_n q_n}{i-r}}{\binom{N_n}{i-1}} \cdot \frac{N_n p_n - r + 1}{N_n - i + 1} = \frac{r}{N_n} \cdot \frac{\binom{N_n p_n}{r}\binom{N_n q_n}{i-r}}{\binom{N_n - 1}{i-1}}.$$

Attesa la doppia corrispondenza

$$T_{i,n} \rightleftharpoons \frac{i}{n},$$

si può dire che ξ_n è una variabile casuale che può assumere i valori

$$(6)\qquad \frac{r}{n}, \frac{r+1}{n}, \ldots, \frac{i}{n}, \ldots, \frac{N_n q_n + r}{n} \text{ (*)},$$

rispettivamente, con le probabilità

$$(7)\qquad \frac{r}{N_n}\frac{\binom{N_n p_n}{r}}{\binom{N_n - 1}{r-1}}, \frac{r}{N_n}\frac{\binom{N_n p_n}{r}}{\binom{N_n - 1}{r}}, \ldots, \frac{r}{N_n}\frac{\binom{N_n p_n}{r}\binom{N_n q_n}{i-r}}{\binom{N_n - 1}{i-1}}, \ldots, \frac{r}{N_n}\frac{\binom{N_n p_n}{r}}{\binom{N_n - 1}{N_n p_n - 1}},$$

ove, ricordiamolo, è $1 < r < N_n p_n$.

È utile osservare che

$$(8)\qquad \frac{r}{N_n}\frac{\binom{N_n p_n}{r}\binom{N_n q_n}{i-r}}{\binom{N_n - 1}{i-1}} = \frac{\binom{i-1}{r-1}\binom{N_n - i}{N_n p_n - r}}{\binom{N_n}{N_n p_n}}.$$

Una tale uguaglianza è notoria. Comunque, per comodità, ricordiamo che risulta

$$(9)\qquad \frac{r}{N_n}\frac{\binom{N_n p_n}{r}\binom{N_n q_n}{i-r}}{\binom{N_n - 1}{i-1}} = \frac{r}{N_n}\frac{1}{r!}\frac{(i-1)!}{(i-r)!} \cdot$$

$$\cdot \frac{N_n p_n (N_n p_n - 1) \ldots (N_n p_n - r + 1) \cdot N_n q_n (N_n q_n - 1) \ldots (N_n q_n - i + r + 1)}{(N_n - 1)(N_n - 2) \ldots (N_n - i + 1)}$$

(*) È appena il caso di osservare che essendo n fissato ed il processo di estrazioni *esaustivo*, è finito il numero dei valori possibili per ξ_n. Infatti, se nelle prime $N_n q_n$ estrazioni dovesse comparire sempre l'insuccesso, appunto perché si tratta di processo esaustivo, nell'urna U_n rimarrebbero *soli* gli $N_n p_n$ successi iniziali. Ne segue che nelle successive estrazioni si otterrebbero soltanto successi; di guisa che di questi se ne totalizzerebbero allora r dopo r estrazioni successive alle precedenti $N_n q_n$. Perciò nel caso più sfavorevole (cioè, subordinatamente al verificarsi dell'insuccesso nelle prime $N_n q_n$ estrazioni) la totalizzazione di r successi si verificherebbe con certezza alla $(N_n q_n + r)$-ma estrazione.

e

$$\frac{1}{r!}\cdot\frac{(i-1)!}{(i-r)!}=\frac{1}{r}\binom{i-1}{r-1}. \tag{10}$$

Allora, sostituendo quest'ultima espressione nella precedente, e poi moltiplicando e dividendo per $(N_n p_n - r)!$ e successivamente per

$$(N_n - i)\cdot(N_n - i - 1)\cdot\ldots\cdot(N_n - N_n p_n + 1)$$

si arriva subito al risultato.

Quindi si può anche dire che ξ_n è una variabile casuale che assume i valori

$$\frac{r}{n}, \frac{r+1}{n}, \ldots, \frac{i}{n}, \ldots, \frac{N_n q_n + r}{n}, \tag{11}$$

rispettivamente con le probabilità

$$\frac{\binom{N_n - r}{N_n p_n - r}}{\binom{N_n}{N_n p_n}}, \frac{r\binom{N_n - r - 1}{N_n p_n - r}}{\binom{N_n}{N_n p_n}}, \ldots, \frac{\binom{i-1}{r-1}\binom{N_n - i}{N_n p_n - r}}{\binom{N_n}{N_n p_n}}, \ldots, \frac{\binom{N_n q_n + r - 1}{r-1}}{\binom{N_n}{N_n p_n}} \tag{12}$$

che si prestano (di qui, l'interesse accennato) per verificare l'uguaglianza

$$\sum_{i=r}^{N_n q_n + r} \frac{\binom{i-1}{r-1}\binom{N_n - i}{N_n p_n - r}}{\binom{N_n}{N_n p_n}} = 1 \text{ (*)}.$$

Infine, agli effetti di una più significativa qualificazione della variabile discreta ξ_n, è opportuna la seguente altra sua rappresentazione.

(*) Basta infatti applicare la nota formula

$$\binom{m}{k} = \sum_{i=r}^{m-k+r} \binom{i-1}{r-1}\binom{m-i}{k-r},$$

a sua volta deducibile dalla

$$\binom{m}{k} = \sum_{i=1}^{m-k+1} \binom{m-i}{k-1}$$

iterando l'uguaglianza di Stifel

$$\binom{m}{k} = \binom{m-1}{k-1} + \binom{m-1}{k}.$$

Dalle uguaglianze (9) e (10) e dalla

$$(N_n-1)(N_n-2)\dots(N_n-i+1)=$$
$$=(N_n-1)(N_n-2)\dots(N_n-j)\dots(N_n-(i-r-1))\cdot$$
$$\cdot(N_n-(i-r))(N_n-(i-r+1))\dots(N_n-(i-r+h))\dots(N_n-(i-1)),$$

si deduce che la probabilità con la quale la variabile casuale ξ_n può assumere il generico valore

$$\frac{i}{n},\ (r\leq i\leq N_n q_n+r) \tag{13}$$

è

$$\binom{i-1}{r-1}p_n^r q_n^{i-r}\prod_{h=0}^{r-1}\frac{N_n-h/p_n}{N_n-(i-r+h)}\prod_{j=0}^{i-r-1}\frac{N_n-j/q_n}{N_n-j}, \tag{14}$$

con la convenzione che ogni prodotto del tipo $\prod_{\nu}^{\nu}$ sia uguale ad 1 per $\nu\leq 0$.

La (14) è sufficiente ad evidenziare che, nelle nuove ipotesi $\mathfrak{I}_1$), $\mathfrak{I}_2$), $\mathfrak{I}_3$), la ξ_n differisce dalla ξ_n trovata nel lav. cit. in (1) per il fatto che l'insieme dei valori che le sono accessibili è *bloccato* a $(N_n q_n+r)/n$; e che la probabilità che le due variabili assegnano ad un medesimo valore accessibile è, questa volta, moltiplicando per il fattore

$$H(n,i,r)\triangleq\prod_{h=0}^{r-1}\frac{N_n-h/p_n}{N_n-(i-r+h)}\prod_{j=0}^{i-r-1}\frac{N_n-j/q_n}{N_n-j}. \tag{15}$$

Per questi fatti imputabili al metodo di estrazione a *blocco* (anziché bernoulliano) dall'uran U_n, si potrebbe anche dire che la variabile casuale ξ_n, indicante il tratto parziale $T_{i,n}$, $(r\leq i\leq N_n q_n+r)$, nel corso del quale si totalizzano r arrivi (o, come suol dirsi, anche nella nuova metrica temporale, il tempo d'attesa dell'r-mo arrivo) è una *variabile di Pascal* (brevemente, *pascalliana*) *bloccata*. Potrebbe anche dirsi: *condizionata*, perché di fatto lo è dalla informazione che gli eventi

$$\xi_n=T_{N_n q_n+r+1,n};\dots;\xi_n=T_{N_n q_n+r+\nu,n};\dots \tag{16}$$

sono diventati (nelle ipotesi attuali, caratterizzate dal metodo di estrazione a blocco) *eventi impossibili* e quindi necessariamente probabilizzabili con probabilità nulla.

È il caso di segnalare che per $i=r$ la (14) porge

$$p_n^r\prod_{h=0}^{r-1}\frac{N_n-h/p_n}{N_n-h}, \tag{17}$$

che è appunto la $P(\xi_n=r/n)$; o meglio, la $P(\xi_n=T_{r,n})$.

Osserviamo inoltre che sussistono le relazioni

$$(18)\qquad \frac{1-(r-1)/N_n p_n}{1-(i-1)/N_n} < \frac{N_n - b/p_n}{N_n-(i-r+b)} \leq \frac{N_n}{N_n-(i-r)} = \frac{1}{1-(i-r)/N_n},$$
$$(r<i<N_n q_n + r;\ b=0,1,\dots,r-1),$$

e le seguenti altre

$$(19)\qquad \frac{1-(i-r-1)/N_n q_n}{1-(i-r-1)/N_n} < \frac{N_n - j/q_n}{N_n - j} < 1,$$
$$(r+1<i<N_n q_n + r;\ j=0,1,\dots,i-r-1)$$

ove, è bene osservare, che il più piccolo valore consentito ad i non è più r, bensì $r+1$, di guisa che $(i-r-1)\geq 0$.

Di qui seguono le disuguaglianze

$$(20)\qquad \left(\frac{1-(r-1)/N_n p_n}{1-(i-1)/N_n}\right)^r < \prod_{b=0}^{r-1} \frac{N_n - b/p_n}{N_n-(i-r+b)} < \frac{1}{(1-(i-r)/N_n)^r},$$
$$(r<i<N_n q_n + r),$$

$$(21)\qquad \left(\frac{1-(i-r-1)/N_n q_n}{1-(i-r-1)/N_n}\right)^{i-r} < \prod_{j=0}^{i-r-1} \frac{N_n - j/q_n}{N_n - j} < 1,\quad (r+1<i<N_n q_n + r),$$

e quindi le seguenti

$$(22)\qquad \left(\frac{1-(r-1)/N_n p_n}{1-(i-1)/N_n}\right)^r \left(\frac{1-(i-r-1)/N_n q_n}{1-(i-r-1)/N_n}\right) < \prod(n,i,r) < \frac{1}{(1-(i-r)/N_n)^r},$$
$$(r+1<i<N_n q_n + r).$$

Queste ultime permettono di stabilire un confine inferiore e superiore per la probabilità $P(\xi_n = T_{i,n})$, con $(r+1)<i<N_n q_n + r$, come appare dalle disuguaglianze

$$(23)\qquad \binom{i-1}{r-1} p_n^r q_n^{i-r} \left(\frac{1-(r-1)/N_n p_n}{1-(i-1)/N_n}\right)^r \left(\frac{1-(i-r-1)/N_n q_n}{1-(i-r-1)/N_n}\right) <$$
$$< \binom{i-1}{r-1} p_n^r q_n^{i-r} \prod(n,i,r) < \frac{\binom{i-1}{r-1} p_n^r q_n^{i-r}}{(1-(i-r)/N_n)^r}.$$

Per $i=r$, la (17) e le (18) porgono

$$(24)\qquad p_n^r \left(\frac{1-(r-1)/N_n p_n}{1-(r-1)/N_n}\right)^r < P(\xi_n = T_{r,n}) = p_n^r \prod_{b=0}^{r-1} \frac{N_n - b/p_n}{N_n - b} < p_n^r$$

le quali sono significative solo per $r>1$, perché per $r=0$ non dicono altro che $p_n^r < p_n^r < p_n^r$.

3. - Studio della variabile casuale che descrive il numero complessivo di tutti i clienti arrivabili nel corso di un fissato intervallo di tempo

Sia $]0, x]$ il fissato intervallo di tempo in discorso, e sia

$$T_{i(n,x),n} = \left(t \in \mathbf{R}^*, \frac{i(n,x)-1}{n} < t \leq \frac{i(n,x)}{n}\right) \tag{25}$$

il particolare tratto parziale (dell'equipartizione D_n) che contiene l'x fissato, e per il quale, ricordiamolo fin d'ora, sussistono le disuguaglianze:

$$-\frac{1}{n} < x - \frac{i(n,x)}{n} \leq 0 ,$$

e quindi le relazioni

$$\lim_{n \to +\infty} \frac{i(n,x)}{n} = x \neq 0 ; \qquad \lim_{n \to +\infty} i(n,x) = +\infty , \tag{26}$$

la convergenza di $i(n,x)/n$ verso x potendo anche essere oscillante (come è stato rilevato nella nota (*) del loc. cit. in (1)).

Indichiamo con η_n la variabile casuale che indica il numero aleatorio totale dei clienti che, complessivamente, nelle ipotesi $\mathfrak{I}_1$), $\mathfrak{I}_2$), $\mathfrak{I}_3$), possono arrivare nel corso del periodo di tempo $]0, x]$; anzi, attesa la metrica dei tempi, nel periodo $]0, i(n,x)/n]$. Appunto per le ipotesi $\mathfrak{I}_1$), $\mathfrak{I}_2$), $\mathfrak{I}_3$) il numero aleatorio η_n equivale al numero complessivo di successi in un campione di taglia $i(n,x)$ estratto *a blocco* (cioè, col metodo esaustivo) dall'urna U_n. Il più piccolo valore accessibile ad η_n allora non è necessariamente lo zero. È il

$$\max\left(0, i(n,x) - N_n q_n\right) . \tag{27}$$

Infatti, se $i(n,x)$ supera il numero complessivo $N_n q_n$ degli insuccessi, potrebbe darsi il caso che in $N_n q_n$ delle $i(n,x)$ estrazioni da U_n (necessarie per formare il campione) comparissero *tutti* gli $N_n q_n$ insuccessi, di guisa che nelle rimanenti $i(n,x) - N_n q_n$ estrazioni comparirebbero allora necessariamente *solo* successi. Perciò il minimo valore accessibile ad η_n è quello indicato in (27). Anche il valore massimo accessibile ad η_n non è necessariamente $i(n,x)$. È infatti, necessariamente, un numero inferiore ad $i(n,x)$ qualora dovesse essere $i(n,x)$ maggiore del numero complessivo $N_n p_n$ dei successi. Perciò la variabile casuale η_n assume i valori interi

$$\max\left(0, i(n,x) - N_n q_n\right), \ldots, k, \ldots \min\left(i(n,x), N_n p_n\right) , \tag{28}$$

il generico dei quali, appunto perché si estrae col metodo esaustivo, è atten-

dibile con la probabilità

$$(29) \qquad P(\eta_n = k) = \frac{\binom{N_n p_n}{k}\binom{N_n q_n}{i-k}}{\binom{N_n}{i}}, \qquad i = i(n, x).$$

Tale espressione, col medesimo ragionamento con cui la (5) è stata trasformata nella (14), può essere presentata anche nel seguente altro modo

$$(30) \qquad P(\eta_n = k) = \binom{i}{k} p_n^k q_n^{i-k} \prod_{h=0}^{k-1} \frac{N_n - h/p_n}{N_n - (i-k+h)} \cdot \prod_{j=0}^{i-k-1} \frac{N_n - j/q_n}{N_n - j},$$

ove, come al solito, si è scritto, per semplicità, solo i al posto del valore $i(n, x)$ che individua il tratto parziale $T_{i(n,x),n}$ contenente il fissato x.

Ebbene le (28) e (29) qualificano la variabile casuale η_n come una *variabile ipergeometrica*: precisamente, coi noti simboli, risulta

$$\eta_n = I_p(i(n, x), N_n p_n, N_n q_n) ;$$

invece le (28) e (30) presentano la η_n come una variante dell'analoga bernoulliana definita dalle (6) e (7) del loc. cit. in [1]. Vi differisce per il fatto che l'insieme dei valori che le sono accessibili è *bloccato*: inferiormente dall'intero $\max(0, i(n, x) - N_n q_n)$ e superiormente dall'intero $\min(i(n, x), N_n p_n)$; ed inoltre per il fatto che le due variabili assegnano ad un medesimo valore accessibile una probabilità che, nel caso attuale, è quella della bernoulliana moltiplicata per il fattore

$$(31) \qquad \Pi(n, i, k) = \prod_{h=0}^{k-1} \frac{N_n - h/p_n}{N_n - (i-k+h)} \cdot \prod_{j=0}^{i-k-1} \frac{N_n - j/q_n}{N_n - j},$$

con la convenzione che sia sempre uguale ad uno ogni prodotto tipo $\prod_{j=0}^{\nu}$, con $\nu \leq 0$.

Per tali varianti, imputabili sempre al metodo di estrazione a *blocco* (anziché bernoulliano) dall'urna U_n, si potrebbe dire che la ipergeometrica η_n è una *bernoulliana* (o, *binomiale*) *bloccata*, o meglio, *condizionata* dall'informazione *a priori* che l'insieme dei valori accessibili ha un confine inferiore e superiore che possono differire, rispettivamente dallo zero ed $i(n, x)$ della bernoulliana. A proposito di questi confini, è il caso di osservare che, col divergere di n, la differenza

$$(32) \qquad i(n, x) - N_n q_n = N_n \left(\frac{i(n, x)}{n} \frac{1}{N_n/n} - 1 + \frac{\lambda}{n} \right)$$

diverge a $-\infty$, perché $i(n, x)/n$ converge a norma della (26) e N_n/n diverge a causa dell'ordine di infinito assunto per N_n.

Quindi per n abbastanza grande (ed è questo il caso che interessa in pratica)

il confine inferiore della nostra ipergeometrica si può ritenere lo zero come per la bernoulliana del loc. cit. in ([1]). Non si può dire altrettanto però per il confine superiore perché i termini: $i(n, x)$; $N_n p_n$ sono entrambi infiniti, col divergere di n, del primo ordine.

Come nel loc. cit. ([1]), si può rappresentare il processo di arrivi a tempi discreti ipergeometrico con una figura come la seguente, ove sul piano ortogonale all'asse dei tempi, nel punto di ascissa $i(n, x)/n$, è riportato lo spettro della variabile ipergeometrica

$$\eta_n = f_p(i(n, x), N_n p_n, N_n q_n) .$$

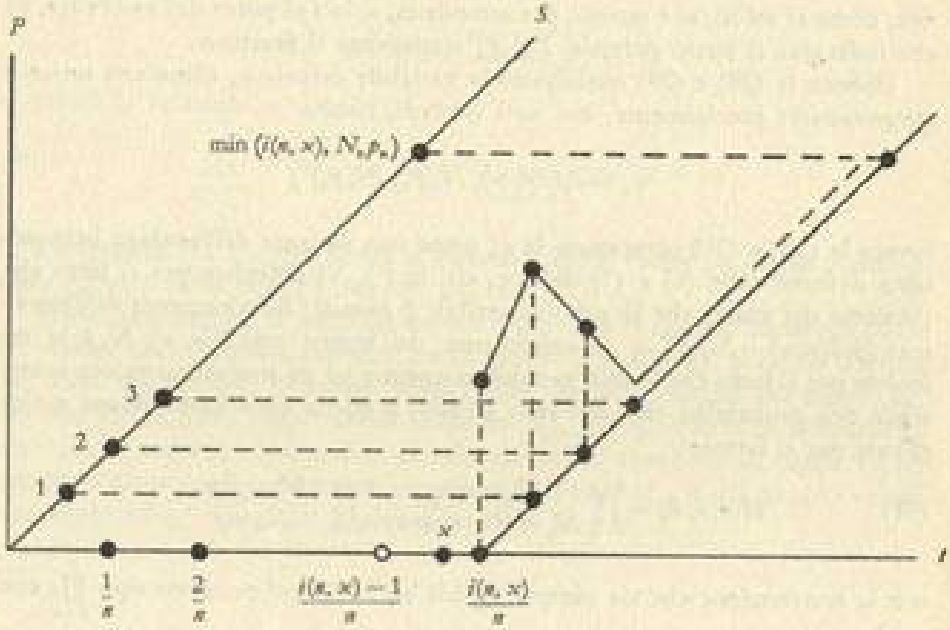

Fig. 1

È appena il caso di ricordare che il valor medio e la varianza di questa variabile sono notoriamente espresse dalle uguaglianze

$$\bar{\eta}_n = \frac{i(n, x)}{n} \lambda ; \qquad \sigma^2_{\eta_n} = \frac{i(n, x)}{n} \lambda \frac{N_n - i(n, x)}{N_n - 1} ;$$

di guisa che, anche questa volta, il numero λ/n appare come numero medio di clienti arrivabili nell'unità di tempo che si è assunta. Con queste premesse è giustificata la definizione posta in chiusura del n. 2.

Rileviamo infine che considerazioni analoghe a quelle svolte in chiusura del precedente n. 3, portano alle disuguaglianze

$$(33)\quad \binom{i(n,x)}{k} p_n^k q_n^{i(n,x)-k} \left(\frac{1-(k-1)/N_n p_n}{1-(i(n,x)-1)/N_n}\right)^k \left(\frac{1-(i(n,x)-k-1)/N_n q_n}{1-(i(n,x)-k-1)/N_n}\right)^{i(n,x)-k} <$$

$$< \binom{i(n,x)}{k} p_n^k q_n^{i(n,x)-k} \prod (n, i(n,x), k) < \frac{\binom{i(n,x)}{k} p_n^k q_n^{i(n,x)-k}}{(1-(i(n,x)-k)/N_n)^k}, \qquad k < i(n,x) \text{ (*)}$$

e, per $K = i(n,x)$

$$(34)\quad p_n^{i(n,x)} \left(\frac{1-(i(n,x)-1)/N_n p_n}{1-(i(n,x)-1)/N_n}\right)^{i(n,x)} < P(\eta_n = i(n,x)) =$$

$$= p_n^{i(n,x)} \prod_{h=0}^{i(n,x)-1} \frac{N_n - h/p_n}{N_n - h} < p_n^{i(n,x)},$$

ove, è bene ricordarlo, per ogni fissato n, $i(n,x)$ è un intero ben determinato.

4. - Studio della variabile casuale che descrive il tempo (aleatorio) separatore di due arrivi successivi

Proveremo che:

condizione necessaria e sufficiente affinché un processo (di arrivi) a tempi discreti sia un processo ipergeometrico $I_p(\lambda/n)$, è che il tempo (aleatorio) separatore di due arrivi successivi sia una variabile casuale τ_n che assume i valori

$$(35)\qquad \frac{1}{n}, \frac{2}{n}, \ldots, \frac{i}{n}, \ldots, \frac{N_n q_n + 1}{n}$$

rispettivamente con le probabilità

$$(36)\quad p_n, p_n q_n \frac{N_n}{N_n - 1}, \ldots, p_n q_n^{i-1} \frac{N_n}{N_n - (i-1)} \prod_{j=0}^{i-2} \frac{N_n - j/q_n}{N_n - j}, \ldots$$

$$\ldots, p_n q_n^{N_n q_n + 1} \frac{N_n}{N_n p_n - 1} \prod_{j=0}^{N_n q_n - 1} \frac{N_n - j/q_n}{N_n - j}$$

salva ancora la convenzione che sia uguale ad 1 ogni prodotto del tipo $\prod_{j=0}^{\nu}$ con $\nu < 1$.

Anche stavolta si potrebbe dire che la variabile casuale τ_n, definita dalle (35) e (36), è una variabile casuale *geometrica bloccata* (o, *condizionata*) dall'in-

(*) A norma delle (22) dovrebbe anche essere $i(n,x) < N_n q_n + k$. Ma questa non è una sostanziale limitazione, perché per quanto si è detto a proposito della (32), la limitazione può ritenersi soddisfatta assumendo n abbastanza grande.

formazione, a priori, che gli eventi

$$(37) \qquad \mathfrak{T}_n = \frac{N_n q_n + 2}{n}, \ldots, \mathfrak{T}_n = \frac{N_n q_n + r}{n} \ldots$$

sono *impossibili*, e quindi necessariamente probabilizzabili con probabilità nulla. Passiamo alla dimostrazione.

Osserviamo intanto che ponendo l'origine dei tempi nell'estremo destro del tratto parziale nel quale è arrivato uno *qualsiasi* dei clienti, il tempo separatore di due arrivi successivi si identifica col tempo d'attesa del primo cliente. Allora la necessità della condizione è implicata dal fatto che nel processo $I_n(\lambda/n)$, caratterizzato con le ipotesi $\mathfrak{I}_1$), $\mathfrak{I}_2$) $\mathfrak{I}_3$), il tempo d'attesa del primo cliente non è altri che la variabile casuale ξ_n, studiata nel numero precedente, relativa al caso $r = 1$. Quindi ponendo $r = 1$ nelle (6) e (7) o meglio nelle (14), si ottengono le (35) e (36). Viceversa, sempre ponendo l'origine dei tempi nell'estremo destro del tratto parziale nel quale è arrivato uno *qualsiasi* dei clienti, la variabile $\mathfrak{T}_n$, descritta in (35) e (36) come tempo separatore di due arrivi successivi in un processo di arrivi, informa che il *primo* cliente potrà arrivare nel generico tratto $T_{i,n}$, $(1 < i \leq N_n q_n + 1)$, con la probabilità

$$p_n q_n^{i-1} \frac{N_n}{N_n - (i-1)} \prod_{j=0}^{i-2} \left(\frac{N_n - j/q_n}{N_n - j} \right),$$

pari a quella con cui, in un campione di taglia i estratto da un'urna come la U_n (associata a D_n), si può verificare un solo successo, e per di più che tale unico successo compaia solo nell'ultima (la i-ma) estrazione, come prescrivono le ipotesi $\mathfrak{I}_1$), $\mathfrak{I}_2$), $\mathfrak{I}_3$); di guisa che il processo è allora ipergeometrico.

È appena il caso di rilevare che la probabilità dell'evento $(\mathfrak{T}_n = i/n)$ ammette le limitazioni evidenziate dalle disuguaglianze

$$(38) \qquad p_n q_n^{i-1} \frac{1}{1-(i-1)/N_n} \left(\frac{1-(i-2)/N_n q_n}{1-(i-2)/N_n} \right)^{i-1} <$$

$$< p_n q_n^{i-1} \frac{N_n}{N_n - (i-1)} \prod_{j=0}^{i-2} \frac{N_n - j/q_n}{N_n - j} < \frac{p_n q_n^{i-1}}{1-(i-1)/N_n},$$

la cui informazione, almeno per $i > 1$, può essere utile nelle applicazioni.

5. - Convergenza del processo di arrivi a tempi discreti ipergeometrico $\mathfrak{I}_n(\lambda/n)$ al processo poissoniano a tempo continuo $\mathfrak{P}(\lambda)$

Seguendo la procedura indicata nel loc. cit. in (1) si tratta di provare nel nostro caso che, col divergere di n, anche

α) la pascalliana *condizionata* ξ_n definita dalle (13) e (14), converge verso la variabile casuale ξ assolutamente continua di Erlang che, in un processo di

arrivi a tempo continuo poissoniano $\mathcal{P}(\lambda)$ con λ numero medio di arrivi nella unità di tempo, descrive l'istante aleatorio nel quale si totalizzano r arrivi. Per $r = 1$ si otterrà che la geometria *condizionata* $\mathcal{G}_n$ convergerà alla variabile $\mathcal{G}$ distribuita con legge esponenziale tipica del tempo separatore di due arrivi successivi in un processo a tempo continuo poissoniano;

β) la variabile casuale ipergeometrica η_n (chiamata bernoulliana *condizionata*) che assume il generico valore naturale k con la probabilità indicata in (29) ed il cui spettro appare in fig. 1 nel piano ortogonale all'asse t nel punto di assissa $i(n, x)/n$, converge verso la poissoniana di valor medio λ.

La tesi α) verrà provata utilizzando la generalizzazione indicata nel n. 4 del loc. cit. in (1) del criterio di compatibilità di Volpato-Bressan (7). Per questo, anche nel nostro caso, prolunghiamo su tutto l'asse reale l'equipartizione D_n di passo $1/n$, e costruiamo un istogramma statistico a colonne, diciamolo H_n, come quello che appare nella figura 2 ove ogni singolo tratto $T_{i,n}$ è sormontato da un rettangolo che lo ammette come base e la cui area, eventualmente nulla, uguaglia la probabilità che la nostra pascalliana *condizionata* ξ_n, assegna all'estremo destro i/n del tratto parziale $T_{i,n}$, implicita essendo

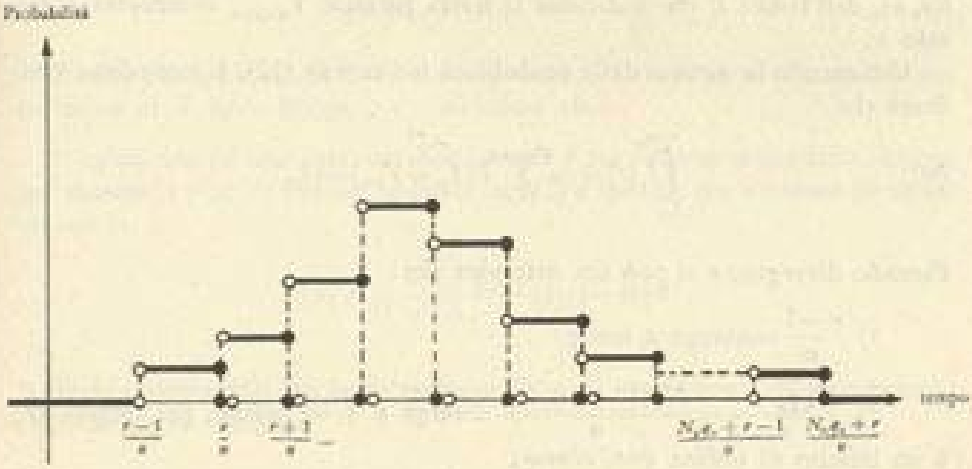

Fig. 2

l'ipotesi che il semiasse reale positivo va considerato non tanto come insieme ordinato di punti ma, come insieme ordinato dei tratti parziali disgiunti della equipartizione D_n che discretizza il tempo.

(7) Cfr. M. Volpato - A. Bressan, *Sulla assoluta continuità di una variabile aleatoria la cui densità è limite di una successione di densità costanti a tratti*, Rend. Sem. Mat. Università di Padova, Vol. 60 (1978), pp. 237-255.

È appena il caso di segnalare che nel nostro caso hanno altezza (e quindi, area) nulla anche tutti i rettangoli che sormontano i tratti parziali successivi al tratto $T_{N_n q_n + r, n}$.

Associata alla variabile discreta ξ_n, anche noi consideriamo la variabile assolutamente continua ξ_n (definita questa volta sull'insieme reale positivo considerato come insieme ordinato di punti) la cui densità, costante a tratti, è la funzione $f_n(x)$ che esprime analiticamente il profilo superiore dell'istogramma. Tale densità è allora definita, per ogni fissato x, dall'uguaglianza

$$(39) \qquad f_n(x) \triangleq \frac{P(\xi_n = T_{i(n,x),n})}{1/n} =$$

$$= \begin{cases} 0, & \text{per } x < \dfrac{r-1}{n}, \\ n\dbinom{i-1}{r-1} p_n^r q_n^{i-r} \prod(n, i, r), & \text{per } \dfrac{r-1}{n} < x < \dfrac{N_n q_n + r}{n}, \\ 0, & \text{per } \dfrac{N_n q_n + r}{n} < x < +\infty, \end{cases}$$

ove per semplicità si è continuato a scrivere i al posto di quel particolare valore $i(n, x)$, dell'indice i, che individua il tratto parziale $T_{i(n,x),n}$ contenente il fissato x.

Utilizzando la somma delle probabilità indicate in (12), è immediato verificare che

$$(40) \qquad \int_{-\infty}^{+\infty} f_n(t)\, dt = \sum_{i=r}^{N_n q_n + r} P(\xi_n = T_{i,n}) = 1 .$$

Facendo divergere n si può ora osservare che:

1) $\dfrac{r-1}{n}$ converge a zero;

2) $\dfrac{N_n q_n + r}{n} = \dfrac{N_n(1 - \lambda/n) + r}{n}$ diverge a $+\infty$ perchè, per ipotesi, N_n è un infinito di ordine due, *almeno*;

3) $\lim\limits_{n \to +\infty} n\dbinom{i-1}{r-1} p_n^r q_n^{i-r} \prod(n, i, r) = \lambda \dfrac{(\lambda x)^{r-1}}{(r-1)!} \exp[-\lambda x]$ (che appunto è la legge di Erlang), perchè come qui di seguito proveremo

$$(41) \qquad \begin{cases} \lim\limits_{n \to +\infty} \prod(n, i, r) = 1 , \\ \text{e} \\ \lim\limits_{n \to +\infty} n\dbinom{i-1}{r-1} p_n^r q_n^{i-r} = \lambda \dfrac{(\lambda x)^{r-1}}{(r-1)!} \exp[-\lambda x] \end{cases}$$

come è stato provato nel n. 5 del loc. cit. in ([1]).

A norma della (15) e (41) basta allora provare che

$$\lim_{n\to+\infty} \prod_{k=0}^{r-1} \frac{N_n - b/p_n}{N_n - (i-r+b)} \prod_{j=0}^{i-r-1} \frac{N_n - j/q_n}{N_n - j} = 1, \tag{42}$$

ove però l'indice i non è più indeterminato, bensì è quel particolare valore $i(n, x)$ per il quale $T_{i(n,x),n}$ contiene l'x fissato, e per il quale quindi valgono le (18). Con questi richiami la prova della (40) è immediata. La fornisce la (22) i cui estremi per l'ordine di infinito assunto per N_n (che, evidentemente, implica la relazione: $\lim_{n\to+\infty} N_n/n = +\infty$) e per le (18) convergono entrambi ad 1. Quindi

$$\lim_{n\to+\infty} f_n(x) \triangleq f(x) = \begin{cases} 0, & \text{per } x \leq 0, \\ \lambda \dfrac{(\lambda x)^{r-1}}{(r-1)!} \exp[-\lambda x], & \text{per } x > 0, \end{cases} \tag{43}$$

per la quale come è stato provato nella (20) del loc. cit. in ([1]) sussiste l'uguaglianza

$$\int_{-\infty}^{+\infty} f(x)\, dx = 1, \tag{44}$$

sufficiente per applicare, anche nel nostro caso, il criterio generalizzato di compatibilità di Volpato-Bressan, e concludere che

esiste una ed una sola variabile casuale ξ assolutamente continua, avente per densità la $f(x)$ di Erlang, indicata in (43) e quindi, per funzione di ripartizione la

$$F(x) \triangleq \frac{\lambda^r}{(r-1)!} \int_0^x t^{r-1} \exp[-\lambda t]\, dt,$$

la quale è *compatibile* con la successione $\{H_n\}$ di istogrammi, ciò significando che sussiste la relazione

$$\lim_{n\to+\infty} \left(f_n(x) - \frac{\Delta F(T_{i(n,x),n})}{\text{mis}\, T_{i(n,x),n}} \right) = 0, \qquad \text{per quasi ogni } x \in \mathbf{R}. \tag{45}$$

Come volevamo la variabile discreta condizionata ξ_n di Pascal, così come la variabile assolutamente continua ξ_n che con l'istogramma le abbiamo associato, converge allora *in legge* verso la variabile ξ di Erlang che, nel processo a tempo continuo poissoniano $\mathcal{P}(\lambda)$, descrive l'istante aleatorio nel quale si totalizzano r arrivi.

Analogo risultato ponendo $r = 1$ per la variabile discreta $\mathfrak{T}_n$ che descrive il tempo separatore di due arrivi successivi nel nostro processo ipergeometrico $\mathfrak{I}_n(\lambda/n)$. Essa converge verso la variabile assolutamente continua $\mathfrak{T}$ distri-

buita con la legge esponenziale

$$g(x) = \begin{cases} 0, & \text{per } x \leq 0, \\ \lambda \exp[-\lambda x], & \text{per } x \geq 0, \end{cases} \tag{46}$$

caratteristica per i tempi separatori di due arrivi successivi in un processo stocastico poissoniano $\mathcal{P}(\lambda)$.

Rimane allora da giustificare la tesi β). Ma ciò è immediato perché, nella ultima parte del n. 5 del loc. cit. in (1) è già stato provato che la bernoulliana assumente i valori

$$k, \qquad 0 < k < i(n, x), \tag{47}$$

rispettivamente con probabilità (usando i simboli attuali)

$$\binom{i}{k} p_n^k q_n^{i-k}, \qquad (0 < k < i(n, x)) \tag{48}$$

converge alla variabile poissoniana

$$\eta = \left\{k;\ P(\eta = k) = \frac{(\lambda x)^k}{k!} \exp[-\lambda x];\ (k = 0, 1, \ldots)\right\} \tag{49}$$

mentre, con l'identico ragionamento dianzi utilizzato si giustifica anche questa volta la relazione

$$\lim_{n \to +\infty} \Pi(n, i, k) = 1.$$

Ne segue che l'ipergeometrica (o, bernoulliana condizionata)

$$\eta_n = \left\{k;\ P(\eta_n = k) = \binom{i}{k} p_n^k q_n^{i-k} \Pi(n, i, k);\ \max(0, i(n, x) - N_n q_n) < \right.$$
$$\left. < k < \min(i(n, x), N_n q_n)\right\} \tag{50}$$

converge alla poissoniana indicata in (39), perché come è già stato osservato

$$\lim_{n \to +\infty} \max(0, i(n, x) - N_n q_n) = 0,$$

e

$$\lim_{n \to +\infty} \min(i(n, x), N_n q_n) = +\infty.$$

Come nel caso studiato in (1), rileviamo che anche qui sussistono le relazioni

$$\lim_{n \to +\infty} \bar{\eta}_n = \bar{\eta} = \lambda x; \qquad \lim_{n \to +\infty} \sigma_{\eta_n}^2 = \sigma_\eta^2 = \lambda x,$$

che evidenziano per λ il significato di numero medio di arrivi nell'unità di tempo.

6. - Un algoritmo per la simulazione delle variabili marginali discrete tipiche del processo a tempi discreti ipergeometrico con le quali si possono approssimare le analoghe leggi nel processo a tempo continuo di Poisson

Atteso il presumibile interesse applicativo che possono avere i risultati qui indicati (soprattutto per il fatto che le nuove ipotesi 3_1), 3_2), 3_3), non richiedono l'indipendenza degli arrivi) è opportuno indicare anche questa volta un diagramma di flusso, sia pur a grandi blocchi, che permetta ad un normale programmatore di organizzare un algoritmo, per la simulazione, col metodo Montecarlo, delle variabili casuali marginali tipiche del processo a tempi discreti ipergeometrico. L'algoritmo col quale simulare la *poissoniana condizionata* appare dal diagramma che segue ove fra i dati compare l'urna U_n da riguardarsi come una (possibile, qualsivoglia) urna la quale contiene inizialmente un insieme formato da N_n numeri equidistribuiti sull'intervallo (0, 1). Ne segue che $N_n p_n$, ($p_n = \lambda/n$), di tali numeri appartengono al sottointervallo $(0, \lambda/n)$ di (0, 1) formando così quel sottoinsieme che abbiamo chiamato *classe dei successi*; e la parte restante $N_n q_n$ appartiene al sottointervallo $(\lambda/n, 1)$ di (0, 1) formando allora quel sottoinsieme che abbiamo chiamato *classe degli insuccessi*. Per quest'urna U_n va ipotizzata l'equicomparibilità dei numeri che vi sono contenuti inizialmente, nonché di quelli che vi rimangono ad ogni stadio di un processo di estrazioni senza reimbussolamento. È pure il caso di osservare che il numero naturale r è un dato che per essere compatibile deve soddisfare la disuguaglianza $r < N_n p_n$, implicante ovviamente la seguente altra $N_n q_n + r < N_n$.

È appena il caso di osservare che al numero i delle estrazioni (determinante l'estremo destro i/n del tratto parziale $T_{i,n}$ nel corso del quale si totalizzano r arrivi) il flow impone *solo implicitamente* la limitazione superiore $N_n q_n + r$.

Infatti, se nelle prime $N_n q_n$ estrazioni la prima frase interrogativa del flow riceve risposta negativa, per il modo con cui è stata composta l'urna U_n, a partire dalla successiva $(N_n q_n + 1)$-ma estrazione quella medesima frase riceverà risposta positiva, e continuerà con tale tipo di risposta fino a totalizzare r risposte positive, consentite peraltro dalla disuguaglianza di compatibilità dei dati: $r < N_n p_n$.

Ne segue che il numero i delle estrazioni certamente non supererà allora il confine superiore $N_n q_n + r$.

Passiamo ora all'algoritmo per la simulazione della variabile ipergeometrica η_n, da noi chiamata anche bernoulliana bloccata (o, condizionata). Anche questa volta fra i dati figura l'urna U_n composta nella identica maniera di quella che compare nel flow precedente.

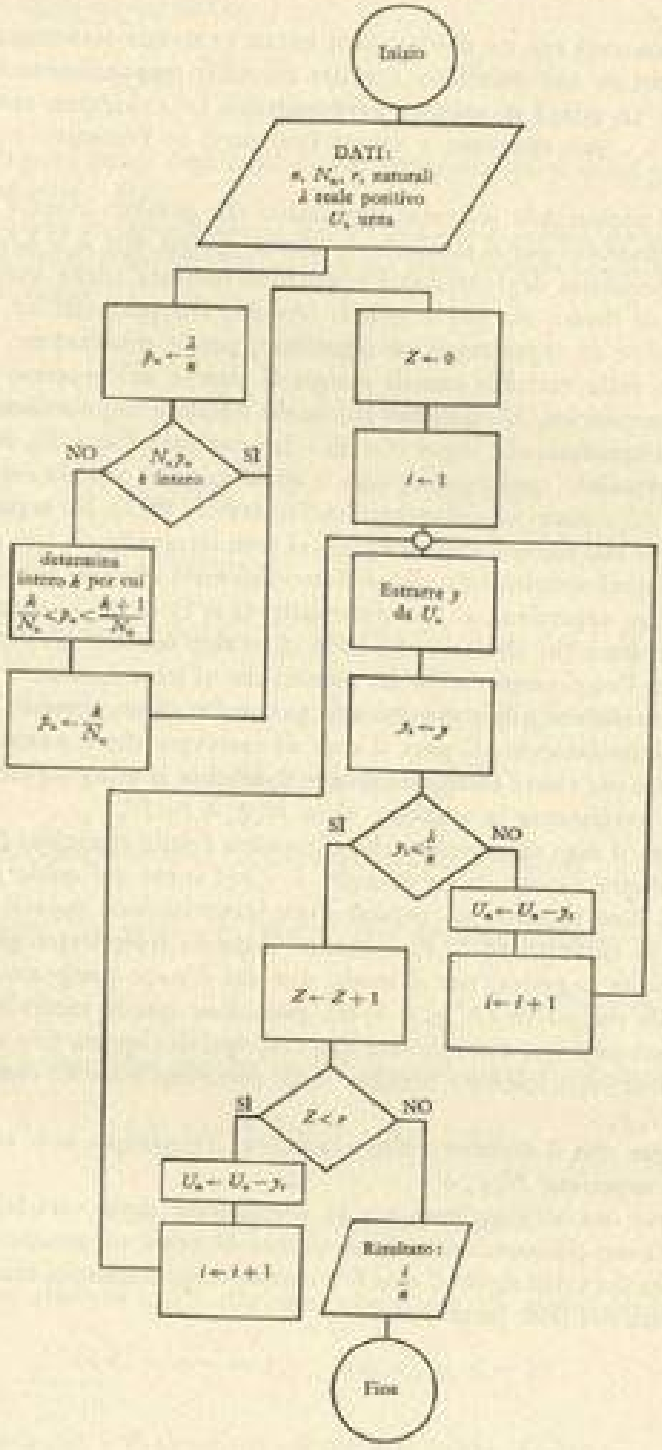
Inizio
DATI:
n, N_n, r, naturali
λ reale positivo
U_n urna
$p_n \leftarrow \frac{\lambda}{n}$
$N_n p_n$ è intero
NO
SI
determina intero k per cui $\frac{k}{N_n} < p_n < \frac{k+1}{N_n}$
$p_n \leftarrow \frac{k}{N_n}$
$Z \leftarrow 0$
$i \leftarrow 1$
Estrarre y da U_n
$y_i \leftarrow y$
$y_i < \frac{\lambda}{n}$
SI
NO
$U_n \leftarrow U_n - y_i$
$i \leftarrow i+1$
$Z \leftarrow Z+1$
$Z < r$
SI
NO
$U_n \leftarrow U_n - y_i$
$i \leftarrow i+1$
Risultato: $\frac{i}{n}$
Fine

Fig. 3

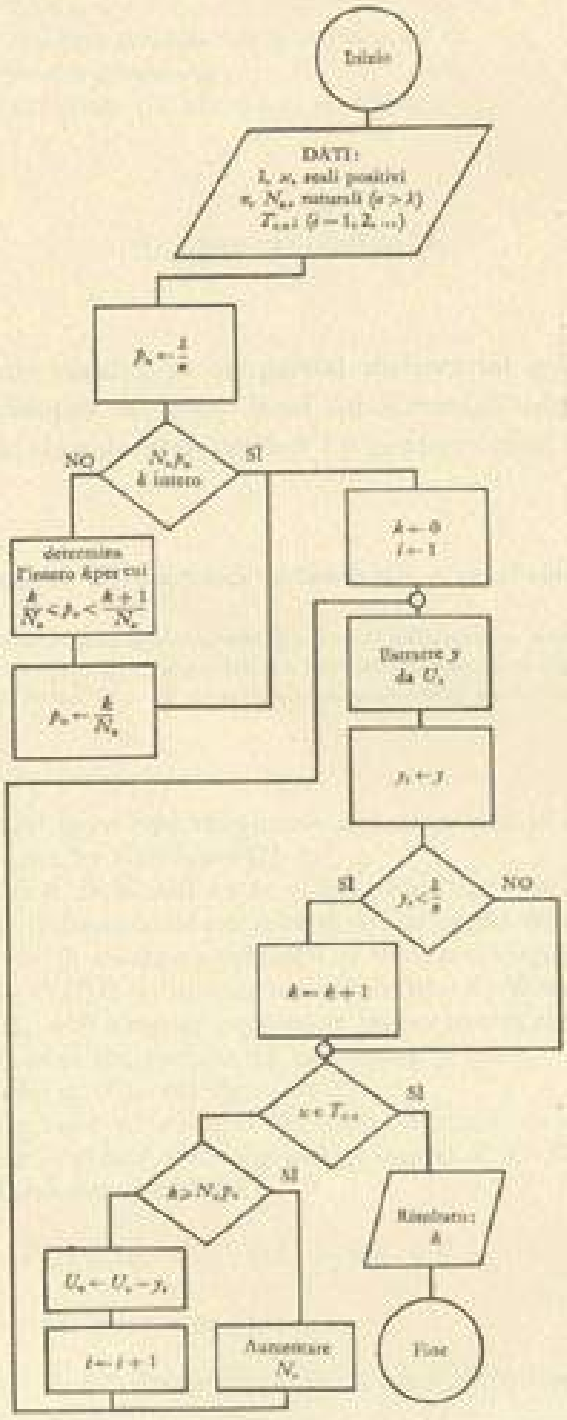
Inizio
DATI:
λ, α, reali positivi
n, N_n, naturali ($n > \lambda$)
$T_{n,i}$ ($i = 1, 2, \ldots$)
$p_n \leftarrow \frac{\lambda}{n}$
$N_n p_n$ è intero
NO
SI
determina l'intero k per cui $\frac{k}{N_n} < p_n < \frac{k+1}{N_n}$
$p_n \leftarrow \frac{k}{N_n}$
$k \leftarrow 0$
$i \leftarrow 1$
Estrarre y da U_n
$y_i \leftarrow y$
$y_i < \frac{\lambda}{n}$
SI
NO
$k \leftarrow k + 1$
$n \in T_{n,i}$
SI
$k \geq N_n p_n$
SI
Risultato: k
$U_n \leftarrow U_n - y_i$
$i \leftarrow i + 1$
Aumentare N_n
Fine

Fig. 4