

疾病統計並に事故統計への彷徨過程の應用 (Ove Lundberg)

丸山, 儀四郎
九州帝國大學理學部

<https://hdl.handle.net/2324/12902>

出版情報 : 統計数理研究. 1 (2), pp.141-155, 1942-03-15. Research Association of Statistical Sciences

バージョン :

権利関係 :

論 著 紹 介

疾病統計並に事故統計への彷徨過程 の應用 (Ove Lundberg)

會員 丸 山 儀 四 郎 (九州帝大理學部)

原著は On random processes and their application to sickness and accident Statistics, University of Stockholm thesis, Upsalla 1940, pp. 1-170. である。著者 O. Lundberg は W. Feller, H. Wold 等と共に Stockholm 大學の H. Cramér を部長とする Department of Mathematical statistics の研究員である。此の論文に於て著者は A. Kolmogoroff [1], W. Feller [1] の彷徨過程 (random process) に關する研究を保險數學方面に應用してゐる。

§1. 緒 論

壺の中に白球, 黒球夫々 Np ケ及び $N(1-p)$ ケある。その中から h 回球をとり出すとし, 各回毎球を戻すとすれば, h 回の中 n 回白球のとり出される確率は, 各試行が獨立であるとすれば

$$(1) \quad \binom{h}{n} p^n (1-p)^{h-n}$$

に依て與へられる。

次に同じ壺から h 回球をとり出すのであるが, 今度は各回毎とり出された球と同じ色の球を $1+N\beta$ ケ戻すとすれば, n 回白球のとり出される確率は次の Pólya-Eggenberger の分布である。

$$(2) \quad \binom{h}{n} \frac{p(p+\beta) \cdots (p+(n-1)\beta) q(q+\beta) \cdots (q+(h-n-1)\beta)}{1(1+\beta) \cdots (1+(h-1)\beta)}, p+q=1.$$

この分布の平均値及び標準偏差の自乗は夫々

$$(3) \quad hp, \quad hpq \frac{1+h\beta}{1+\beta}$$

となる。而して h 回の球の取出しを行ふ途中 m 回目迄に n 回白球がとり出された事を知つた場合次の $m+1$ 回目のとり出しに於て白球の取り出される (條件付) 確率は

$$(4) \quad \frac{Np + \mu N\beta}{N + mN\beta} = \frac{p + \mu\beta}{1 + m\beta}$$

であつて, m 回迄にどの様な順序で白球がとり出されたかと云ふ事には無關係である。此處に擧げた例は球をとり出すと言ふ彷徨過程であるが, 一般に一つの彷徨過程に於て, 途中の或る段階 (或る時刻) 迄に彷徨過程が實現された全経過を知つてゐたとしても, その後の経過を確率論的に記述するためには單に該段階に於ける状態のみで充分である場合, かゝる彷徨過程の特性を influence globale (Pólya) 又は Stochastisch-definit (Kolmogoroff) と呼ぶ。以下かゝる特性を有する彷徨過程のみを取扱ふ。

$p=t/h, \beta=bt/h$ と置き $h \rightarrow \infty$ にすれば (1), (2), (3) は夫々

$$(1)' \quad P_n(t) = \frac{t^n}{n!} e^{-t} \quad (n=0, 1, \dots)$$

$$(2)' \quad P_n(t) = \left(\frac{t}{1+bt} \right)^n \cdot \frac{1 \cdot (1+b) \cdots (1+(n-1)b)}{n!} (1+bt)^{-\frac{1}{b}} \quad (n=0, 1, \dots)$$

$$(3)' \quad t, t(1+bt)$$

となる。(1)' は Poisson 分布, (2)' は (1)' に傳播 (Contagion) の概念を入れる事に依て得られた Pólya-Eggenberger の分布で, t を時間の parameter と考へれば (1)', (2)' なる分布函数を持つ確率變數 n_t に依て彷徨過程が定義される。 $p=t/h, mt/h=\tau, t/h=\Delta t$ と置けば (4) は

$$(4)' \quad \frac{1+\mu b}{1+\tau b} \Delta t$$

となる。之は τ なる時刻までに μ 回考へる事象が現はれた事を知つた場合, 次の Δt 時間内に該事象が一回現はれる確率に漸近的に等しいと考へる事が出来る。

之に反し Greenwood-Yule [1] は, 全く傳播を考へずに (2)' を導いてゐる。即ち Pearson の Type III の分布

$$(5) \quad w(x; c, r) = cr \frac{r^{r-1}}{\Gamma(r)} c^{-cx} \quad (x > 0; c > 0, r > 0)$$

に於いて $c=1/b, r=1/b$ と置いて得られる分布

$$(6) \quad w_b(x) = w\left(x; \frac{1}{b}, \frac{1}{b}\right) = \frac{1}{b} \cdot \frac{\left(\frac{x}{b}\right)^{1/b-1}}{\Gamma\left(\frac{1}{b}\right)} e^{-\frac{x}{b}}$$

に依て積分

$$(7) \quad \int_0^\infty \frac{(tx)^n}{n!} e^{-tx} w_b(x) dx$$

を求めれば丁度これが (2)' になる。即ち x なる parameter を持ち, 確率分布が時間 t に對して $(tx)^n e^{-tx}/n!$ で與へられる様な彷徨過程の集團を考へ, 各彷徨過程はこの集團に於て (6) に従て分布してゐると考へた時, 集團全體の示す時間的推移としてあたかも傳播的であるかの様に見える。従て 彷徨過程が (2)' に依て與へられるからと云つて, 直に b を以て傳播を表はす常數であると推断する事は出来ない。

§2. 彷徨過程を定義する微分方程式

§1 に取扱つた様な問題を統一的に論ずる方法として Kolmogoroff, Feller の方法がある。それは微分方程式の解として彷徨過程を定義する方法である。

確率變數 x_t は任意に與へられた次元の Euclid 空間 R を動くと考へ, x_t が $t=t_0$ に於て x なる値を實現した事を知つた場合, 時刻 t に於て x_t が考へる空間の Borel 集合 A に入つてゐる確率を $P(t_0, x; t, A)$ とし又 $P(t, A) = \Pr(x_t \in A)$ とする。前者は à postériori 確率, 後者は à priori 確率である。 P が確率である以上次の諸性質を満足しなければならない。

$$(P) \quad P(s, x; t, A) \geq 0$$

$$(T) \quad \int_R dP(s, x; t, A) = 1$$

$$(F) \quad P(s, x; t, A) = \int_{\xi \in R} P(\tau, \xi; t, A) dP(s, x; \tau, X), \quad s < \tau < t.$$

(Kolmogoroff-Chapman 微分方程式)

但し勿論上の積分が考へられるために $P(s, x; t, A)$ が x に関して例へば B -measurable であるとする。又 $P(s, x; t, A)$ が s, t に関して連続であるといふ自然な假定を設ける。即ち

$$(I) \quad \lim_{t \rightarrow s} P(s, x; t, A) = \varepsilon(x, A),$$

$$(I^*) \quad \lim_{t \rightarrow s} P(s, x; t, A) = \varepsilon(x, A).$$

但し $\varepsilon(x, A)$ は集合 A の characteristic function. 今後確率とは常に (P), (T), (F), (I), (I*) を満足するものとする。process を continuous 及 discontinuous process に分ける。 Σ_δ を x 中心とし δ なる半径の球とする

$$P(t, x; t + \Delta t, R - \Sigma_\delta) = \int_{R - \Sigma_\delta} dP(t, x; t + \Delta t, A) = o(\Delta t)$$

を満足する彷徨過程を continuous process と云ふ。而して上の式に於て右邊を $o(\Delta t)$ とした式を満足する彷徨過程を discontinuous process と云ふ。之を精確に表現するために次に二つの函数を導入する。 x_t が t に於て x なる値をとつた事を知つた場合に、次の Δt 時間内に random change を受ける確率を、 Δt が充分小なる時漸近的に $p(t, x) \Delta t$ とし、次に random change が起きた事を知つた場合に $x_{t+\Delta t} \in A$ となる確率を、 Δt が充分小なるとき漸近的に $\Pi(t, x, A)$ とする。ここに $p(t, x)$ は §1 の (4)' の Δt の係数に對應する量で之を intensity function と呼ぶ。然るとき discontinuous process は

$$(8) \quad \begin{aligned} P(t, x; t + \Delta t, A) &= \\ &= [1 - p(t, x) \Delta t] \varepsilon(x, A) + \Delta t p(t, x) \Pi(t, x, A) + o(\Delta t) \end{aligned}$$

に依つて定義される。今後は専ら discontinuous process を考へる。今之が満足する基礎方程式を導く。(F) と (8) に依り

$$\begin{aligned} P(s, x; t + \Delta t, A) &= \int_{\xi \in R} P(t, \xi; t + \Delta t, A) dP(s, x; t, X) \\ &= \int_{\xi \in R} [1 - p(t, \xi) \Delta t] \varepsilon(\xi, A) dP(s, x; t, X) + \Delta t \cdot \int_{\xi \in R} p(t, \xi) \\ &\quad \times \Pi(t, \xi, A) \cdot dP(s, x; t, X) + o(\Delta t) = P(s, x; t, A) - \\ &\quad - \Delta t \int_{\xi \in A} p(t, \xi) dP(s, x; t, X) + \Delta t \int_{\xi \in R} p(t, \xi) \Pi(t, \xi, A) dP(s, x; t, X) + o(\Delta t). \end{aligned}$$

故に

$$(9) \quad \begin{aligned} \frac{\partial}{\partial t} P(s, x; t, A) &= - \int_A p(t, \xi) dP(s, x; t, X) \\ &\quad + \int_{\xi \in R} p(t, \xi) \Pi(t, \xi, A) dP(s, x; t, X). \end{aligned}$$

同様に (F) の右邊の $P(s, x; \tau, X)$ に (8) を代入して

$$(9)^* \quad \frac{\partial}{\partial s} P(s, x; t, A) = p(s, x) [P(s, x; t, A) - \int_{x \in R} P(s, \xi; t, A) d\Pi(s, x, X)]$$

(9), (9)* が discontinuous process に対する基礎方程式である。

今 x_i の取り得る状態が可附番簡であるとすれば彷徨過程は Markoff 過程になる。この状態に番號をつける事とすれば $P(s, x; t, A)$, $p(t, x)$, $\Pi(t, x, A)$ に相當しては, m, n を自然數として

$$P_{m,n}(s, t) \equiv P(s, m; t, n), \quad p_n(t) \equiv p(t, n), \quad \Pi_{m,n}(t) \equiv \Pi(t, m, n)$$

を考へればよい。之に對して (P), (T), (F), (I), (I*) を書けば, 積分を和で置換へる事に依り

$$(P_1) \quad P_{m,n}(s, t) \geq 0,$$

$$(T_1) \quad \sum_{n=0}^{\infty} P_{m,n}(s, t) = 1,$$

$$(F_1) \quad P_{m,n}(s, t) = \sum_{v=0}^{\infty} P_{m,v}(s, \tau) P_{v,n}(\tau, t), \quad s < \tau < t,$$

$$(I_1) \quad \lim_{t \rightarrow s} P_{m,n}(s, t) = \epsilon_{m,n},$$

$$(I_1)^* \quad \lim_{s \rightarrow t} P_{m,n}(s, t) = \epsilon_{m,n}.$$

こゝに $\epsilon_{m,n}$ は Kronecker の記號である。即ち $\epsilon_{m,n} = 0 (m \neq n)$, $\epsilon_{m,n} = 1 (m = n)$ 。而してこの場合 (9), (9)* に相當する基礎方程式は

$$(10) \quad \frac{\partial}{\partial t} P_{m,n}(s, t) = -p_n(t) P_{m,n}(s, t) + \sum_{v=0}^{\infty} p_v(t) \Pi_{v,n}(t) P_{m,v}(s, t)$$

$$(10)^* \quad \frac{\partial}{\partial s} P_{m,n}(s, t) = p_m(s) \left[P_{m,n}(s, t) - \sum_{v=0}^{\infty} P_{v,n}(s, t) \Pi_{m,v}(s) \right].$$

この特別な場合として重要なのは, $\Pi_{m,n}(t)$ が特に

$$(11) \quad \Pi_{m,n}(t) = \epsilon_{m,n-1} = \begin{cases} 1 & n = m+1 \\ 0 & n \neq m+1 \end{cases}$$

を満足する場合である。この場合, $\Pi_{m,n}(t)$ の定義から分る様に或る瞬間に x_i が m と云ふ状態にあれば, 次に x_i が random change を受けて取り得る状態は $m+1$ である。この様な彷徨過程を elementary process と呼ぶ。この彷徨過程は或種の繁殖とか崩壊 (disintegration) 現象の確率論的な表現と考へる事が出来やう。此の論文で取扱ふのは elementary process に限る。之に對する式は (10), (10)* から基礎方程

$$(12) \quad \frac{\partial}{\partial t} P_{m,n}(s, t) = -p_n(t) P_{m,n}(s, t) + p_{n-1}(t) P_{m,n-1}(s, t)$$

$$(12)^* \quad \frac{\partial}{\partial s} P_{m,n}(s, t) = p_m(s) [P_{m,n}(s, t) - P_{m+1,n}(s, t)].$$

となる。而して discontinuous process を定義した (8) はこの場合特に

$$(13) \quad P_{m,n}(t, t+dt) = \begin{cases} 1 - p_m(t) dt + o(dt) & \dots n = m \\ p_m(t) dt + o(dt) & \dots n = m+1 \\ o(dt) & \dots n = m+2 \end{cases}$$

となる。吾々は (12), (12)* に於て $p_i(t) \geq 0$ を與へられた連続函數とし之を夫々條件 (I₁), (I₁)* の

下に解いて elementary process を求めるのであるが、その場合、解 $P_{m,n}(s, t)$ は (P_1) , (T_1) , (F_1) 及び (13) を満足しなければならない。

W. Feller [1] に依れば $p_n(t)$ が区間 $T_1 \leqq t \leqq T_2$ に於て、連続で一様に有界 $p_n(t) \leqq K(T_1, T_2)$ であれば (12) (12)* は (I_1) , $(I_1)^*$ の下に同一の連続な単一解 $P_{m,n}(s, t)$ を有し、 $s, t, t+dt$ がこの区間に属すれば $P_{m,n}(s, t)$ は所要の条件 (P_1) , (T_1) , (F_1) 及び (13) を満足する。

著者はこの W. Feller の定理を少しく改良した形に於て証明してゐる。但し elementary process に対しては条件 (11) から $P_{m,n}(s, t) = 0 \quad n < m$, 従て方程式 (12), (12)* に於て $n \geqq m$ の場合を考へればよい。然るとき (12), (12)* は次の形をとる。

$$(D) \quad \begin{cases} \frac{\partial}{\partial t} u_{m,n}(s, t) = -p_n(t)u_{m,n}(s, t) + p_{n-1}u_{m,n-1}(s, t) \\ \frac{\partial}{\partial t} u_{n,n}(s, t) = -p_n(t)u_{n,n}(s, t) \end{cases}$$

$$(D)^* \quad \begin{cases} \frac{\partial}{\partial s} u_{m,n}^*(s, t) = p_m(s) [u_{m,n}^*(s, t) - u_{m+1,n}^*(s, t)] \\ \frac{\partial}{\partial s} u_{n,n}^*(s, t) = p_n(s)u_{n,n}^*(s, t). \end{cases}$$

但し $P_{m,n}$ の代りに $u_{m,n}$ と書いたのはこの解が一般には確率としての所要の条件を満足しないからである。然るとき次の定理が証明される。

定理 $p_n(t) \geqq 0$ を連続函数とすれば (D), (D)* は単一な解 $u_{m,n}(s, t)$, $u_{m,n}^*(s, t)$ を有し次の性質満足をする。

$$(a) \quad u_{m,n}(s, t) = u_{m,n}^*(s, t)$$

$$(b) \quad u_{m,n}(s, t) = \sum_{v=m}^n u_{m,v}(s, t) u_{v,n}(s, t) \quad (s \leqq \tau \leqq t)$$

$$(c) \quad u_{m,n}(\tau, \tau + dt) = \begin{cases} 1 - p_m(\tau)dt + o(dt) \cdots n = m \\ p_m(\tau)dt + o(dt) \cdots n = m + 1 \\ o(dt) \cdots n \geqq m + 2 \end{cases}$$

$$(d) \quad u_{m,n}(s, t) \geqq 0$$

$$(e) \quad \lim_{n \rightarrow \infty} \sum_{v=m}^n u_{m,v}(s, t) \leqq 1, \quad \lim_{n \rightarrow \infty} \int_s^t p_n(\tau) u_{m,n}(s, \tau) d\tau > 1.$$

尙、若し更に

$$(14) \quad \sum_{n=m}^{\infty} \frac{1}{\text{Max}_{s \leqq \tau \leqq t} p_n(\tau)} = +\infty$$

ならば

$$(f_1) \quad \sum_{n=m}^{\infty} u_{m,n}(s, t) = 1.$$

之に反し、若し

$$(15) \quad \sum_{n=m}^{\infty} \text{Min}_{s \leqq \tau \leqq t_0} p_n(\tau) < t_0 - s$$

ならば $t' < t_0$ なるが t' 存在して $t > t'$ なる如きすべての t に對して

$$(f_2) \quad \sum_{n=m}^{\infty} u_{m,n}(s, t) < 1.$$

即ち (14) は $u_{m,n}(s, t)$ が確率として所要の条件 $(P_1), (T_1), (F_1)$ 及び (13) を満足するために十分である。

次に時間の尺度変換と云ふ事を考へる。今 $\lambda = \lambda(t)$ を正の連続な導函数を有する函数とする。t から λ へ變數變換を行へば $(D), (D)^*$ から容易に分る様に intensity function $p_n(t)$ は變換されて $p_n(t) \frac{dt(\lambda)}{d\lambda}$ となる。但し $t=t(\lambda)$ は $\lambda=\lambda(t)$ の逆函数。應用上重要な尺度は平均値 $v_n(t)$ に依る

$$(16) \quad \lambda = \lambda(t) \equiv v_1(t) \sum_{n=1}^{\infty} n P_n(t),$$

及び

$$(17) \quad \lambda = \int_0^t p_0(\tau) d\tau = -\log P_0(t)$$

で、前者に依る尺度を Operational time scale と云ふ。

$(D), (D)^*$ の解で確率としての条件を満足するもの、即ち $P_{m,n}(s, t)$ が

$$P_{m,n}(s, t) = P_{m,n}(0, t-s)$$

を満足するとき彷徨過程は齊時的 (time homogeneous) であると云ふ。若し elementary process の intensity function が $p_n(t) = p_n \cdot v(t)$ なる形に書けたとすれば (17) に依て尺度を λ に變換すれば intensity function は

$$(18) \quad \bar{p}_n(\lambda) = p_n \cdot v(t) \cdot \frac{1}{p_0 v(t)} = \frac{p_n}{p_0}$$

となる。而して $(D), (D)^*$ の解を explicit に書いて歸納法で證明すれば容易に分る様に、彷徨過程が齊時なるために必要にして且つ十分なる条件は $p_n(t)$ が t に無關係なる事である。従て $p_n(t) = p_n \cdot v(t)$ なるときは (17) に依り λ に尺度を變換すれば λ なる尺度に於ては彷徨過程は齊時になる。

§3. 齊時彷徨過程

$p_n(t)$ は時間に無關係、之を p_n とすれば §2 の定理から容易に分る事は $(D), (D)^*$ の解が所要の条件を満足するために必要且つ十分な条件はすべての m に對して

$$(19) \quad \sum_{n=m}^{\infty} \frac{1}{p_n} = +\infty$$

なる事である。然るとき $(D), (D)^*$ の解の中 $P_{0,m}(0, t)$, 即ち (彷徨過程は始め $m=0$ なる状態から出發したと標準化する事に依り) a priori 確率 $P_{0,m}(t) \equiv P_{0,m}(0, t)$ は實際 $e^{-p_0 t} (v=0, 1, \dots)$ の多項式で書下す事が出来る。即ち

$$(20) \quad P_n(t) = p_0 p_1 \cdots p_{n-1} \sum_{j=0}^n \frac{e^{-p_0 t}}{\prod_{v=j}^{n-1} (p_v - p_j)} \sum_{j=0}^n \frac{p_j}{p_n} \frac{D_n^{(j)}}{D_n} e^{-p_0 t}$$

但し D_n は Vandermonde の行列式

$$D_n = \begin{vmatrix} 1 & p_0 & p_0^2 & \cdots & p_0^n \\ 1 & p_1 & p_1^2 & \cdots & p_1^n \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & p_n & p_n^2 & \cdots & p_n^n \end{vmatrix},$$

$D_n^{(j)}$ は D_n の第 1 列 (column), 第 j 行 (row) の餘因子。

或は

$$(21) \quad P_n(t) = p_0 p_1 \cdots p_{n-1} t^n e^{-tp_0} \int_0^1 ds_1 \cdots \int_0^{s_{n-2}} ds_{n-1} \cdot \int_0^{s_{n-1}} ds_n e^{-t \sum_{i=1}^n s_i (p_i - p_{i-1})}$$

§1の(2)'に依て定義される彷徨過程を Pólya process と云ふ。之は (D), (D)* から容易に分る様に

$$(23) \quad p_n(t) = \frac{1+bn}{1+bt}$$

を intensity function とする elementary process である。従て §2 の (17), 即ち $\lambda = -\log P_0(t) = \frac{1}{b}(1+bt)$ で尺度變換をすれば, λ に就いて彷徨過程は齊時になり intensity function は $p_n = 1+bn$, à priori 確率は (21) から

$$(24) \quad \bar{P}_n(\lambda) = \prod_{\nu=1}^{n-1} (1+b\nu) \lambda^n e^{-\lambda} \cdot \frac{\left(\int_0^1 e^{-b\lambda s} ds\right)^n}{n!} = \frac{\prod_{\nu=1}^{n-1} (1+b\nu)}{b_n \cdot n!} e^{-\lambda} (1-e^{-b\lambda})^n.$$

他方に於て

$$(25) \quad \bar{p}_1(\lambda) = \sum_1^{\infty} n \bar{P}_n(\lambda) = e^{-\lambda} \sum_{n=1}^{\infty} n \binom{-1/b}{n} (e^{-b\lambda} - 1)^n = \frac{e^{b\lambda} - 1}{b} = t,$$

従て λ から operational time scale に變換すれば, 時間の尺度は元に戻る。

§4. Compound Poisson Process.

Greenwood-Yule の考へを一般化して compound Poisson Process なるものを考へる。それは次の如く定義される。

à priori 確率 $P_n(t)$ が

$$(26) \quad P_n(t) = \int_0^{\infty} \frac{(tx)^n}{n!} e^{-tx} dU(x)$$

に依て與へられる elementary process を compound Poisson process と呼ぶ。但し $U(x)$ は分布函数で $U(+\infty) - U(0) = 1$ 。

この彷徨過程の $p_n(t)$ は (D), (D)* から容易に分る様に

$$(27) \quad p_n(t) = \frac{\int_0^{\infty} x^{n+1} e^{-tx} dU(x)}{\int_0^{\infty} x^n e^{-tx} dU(x)} = - \frac{P_0^{(n+1)}(t)}{P_0^{(n)}(t)}$$

に依て與へられる。 $p_n(t)$ は次の定差方程式を満足する。

$$(28) \quad p_{n+1}(t) = p_n(t) - \frac{p_n'(t)}{p_n(t)} \quad (n=0, 1, \dots).$$

一般に $t > 0$ に對し何回でも微分出來て

$$(29) \quad (-t)^n P^{(n)}(t) \geq 0$$

を満足する函数 $P(t)$ は ($t > 0$ で) completely monotonic であると云ふ。函数が $t > 0$ に於いて completely monotonic なるための必要十分條件は, 有限區間で有界な非減少函数 $G(x)$ を以て函数が Laplace-Stieltjes 積分

$$\int_0^{\infty} e^{-xt} dG(x)$$

で表はされる事である (Bernstein [1], Feller [2]). そうすると Compound Poisson process の intensity function $p_n(t)$ は completely monotonic な函数 $P_0(t)$ を以て

$$(30) \quad p_n(t) = -\frac{P_0^{(n+1)}(t)}{P_0^{(n)}(t)}$$

なる形に書かれるわけであるが、又逆に

$$(1) \quad t > 0 \text{ で } P_0(t) \text{ は completely monotonic,}$$

$$(2) \quad \lim_{t \rightarrow 0} P_0(t) = 1$$

なる $P_0(t)$ を以て、 $p_n(t)$ が (30) で與へられるとすれば $p_n(t)$ は compound Poisson process の intensity function になりその à priori 確率 $P_n(t)$, à postériori 確率 $P_{m,n}(s, t)$ は

$$(31) \quad P_n(t) = \frac{(-t)^n}{n!} P_0^{(n)}(t) = \int_0^\infty \frac{(tx)^n}{n!} e^{-tx} dU(x) \quad (n=0, 1, \dots)$$

$$(32) \quad P_{m,n}(s, t) = (-1)^{n-m} \frac{(t-s)^{n-m}}{(n-m)!} \frac{P_0^{(n)}(t)}{P_0^{(m)}(s)}$$

に依て與へられる事を證明する事が出来る。

又 completely monotonic な函数 $P_0(t)$ をて以最初

$$P_0'(t) = -p_0(t) P_0(t)$$

を以て $p_0(t)$ を與へ、次に (28) に依て順次 $p_1(t), p_2(t), \dots$ を定義してゆけば $p_n(t)$ は (30) の形に書かれ、従てかゝる $p_n(t)$ は compound Poisson process を與へる。故に (28) を compound Poisson process の特性と考へる事が出来る。

次に (32) の右邊を (31) に依て書換へれば

$$(33) \quad P_{m,n}(s, t) = \binom{n}{m} \left(\frac{s}{t}\right)^m \left(1 - \frac{s}{t}\right)^{n-m} \frac{P_n(t)}{P_m(s)} \quad (t > s).$$

今

$$P_m(s) P_{m,n}(s, t) = P_n(t) Q_{m,n}(s, t) \quad (t > s)$$

を以て $Q_{m,n}(s, t)$ を定義すれば、 $Q_{m,n}(s, t)$ は t と云ふ時刻に x_t が n となつた事を知つた場合に時刻 $t=s$ に於て x_s が m であつた確率 (inverse probability) に他ならない。そうすると (33) に依て compound Poisson process の inverse probability は二項分布

$$(34) \quad Q_{m,n}(s, t) = \binom{n}{m} \left(\frac{s}{t}\right)^m \left(1 - \frac{s}{t}\right)^{n-m}$$

になる。逆に inverse probability が二項分布 (34) になり従つて $P_{m,n}(s, t)$ が (33) に依て與へられれば、彷徨過程は compound Poisson process になる事が容易に證明出来る。従つて inverse probability が二項分布であると云ふ事が又 compound Poisson process の特性である。

compound Poisson process を定義する分布函数 $U(x)$ を特に

$$U'(x) = w_b(x) = \frac{1}{b} \cdot \frac{\left(\frac{x}{b}\right)^{1/b-1}}{\Gamma\left(\frac{1}{b}\right)} e^{-\frac{x}{b}}$$

とすれば Pólya process が得られる。その intensity function は (30) から

$$p_n(t) = \frac{1+bn}{1+bt}$$

となり、勿論 (23) と一致する。之を見ると (a) $p_n(t) = p_n \cdot v(t)$, (b) $p_n(t) = \psi_0(t) + n\psi_1(t)$ なる形をしてゐる。然し逆に intensity function が, (a) 又は (b) の形に書ける様な compound Poisson process は Pólya process に限る事が證明出来、従つて (a) 又は (b) は Pólya process の特性である。

§5. random change の大きさを一般にした compound Poisson process.

今迄考へた彷徨過程に於ては $\Pi_{m,n}(t) = \varepsilon_{m,n-1}$ であつた。即ち確率變數 x_t は m から $m+1$ へと變化するのであつて random change の大きさを常に 1 と指定してゐる。此の § に於ては random change の大きさを指定せず唯一定の確立法則に従ふものとし、random change の起る回数は elementary process になつてゐるものとする。又確率變數のとり得る値は實數とする。そうすると確率變數 x_t と共に random change の起る回数を表はす確率變數 n_t を考へ、二次元の彷徨過程として考へるのが便利である。之に對應して一般なる彷徨過程に對して成立つ微分方程式 (9) を考へる場合、集合 A は

$$(35) \quad A = A(\mathbf{n} = n, \mathbf{x} \leq x)$$

と考へればよい。今 t なる時刻迄に random change が n 回起り、従つて n_t が n なる値を取り、 x_t が x なる値を取つた事を知つた場合、intensity function を $p_n(t, x)$ とし、時刻 t の後に x_t の受ける random change の大きさが u を超えない確率を $V(t, n, x, u)$ とする。そうすると t 時刻に (m, x_0) なる状態にあつた確率變數 (n_t, x_t) が random change を受けて状態 (35) になる確率は、 n_t が elementary process になつてゐる事から、 $\Pi = \varepsilon_{m,n-1} V(t, m, x_0, x - x_0)$ で與へられる。之が今考へる場合の微分方程式 (9) に於ける函數 $\Pi(t, x, A)$ に相當する函數である。今後問題を簡單にして $p_n(t, x) = p_n(t)$, $V(t, n, x, u) = V(u)$ と書ける場合を取扱ふ。かゝるときは微分方程式 (9) の解は

$$(36) \quad P_{m,n}(s, t) V_{n-m}(x - x_0)$$

に依て與へられる。但し $V_k(u)$ は $V(u)$ の k 回の convolution, $P_{m,n}(s, t)$ は elementary process n_t の à postériori 確率である。何となれば微分方程式 (9) に (36) を代入すると、此の場合左邊は

$$(37) \quad \frac{\partial}{\partial t} P_{m,n}(s, t) \cdot V_{n-m}(x - x_0),$$

右邊は

$$(38) \quad -p_n(t) \cdot P_{m,n}(s, t) V_{n-m}(x - x_0) + p_{n-1}(t) \cdot P_{m,n-1}(s, t) \int_{-\infty}^{\infty} V(x - \xi) dV_{n-1-m}(\xi - x_0)$$

となり $P_{m,n}(s, t)$ が (12) を満足する事から (37) と (38) は等しくなる。 $(m, x_0) = (0, 0)$ なる状態から彷徨過程が始まつたとすれば彷徨過程の à priori 確率は $P_n(t) V_n(x)$ で與へられ、従て x_t の分布函數は

$$(39) \quad F(x, t) = \sum_{n=0}^{\infty} P_n(t) V_n(x)$$

となる。先に考へた random change の大きさを 1 なる elementary process は $V(x) = \varepsilon(x-1)$ なる

特別の場合になる。そして實際其の場合 (39) は $V_n(x) = \varepsilon(x-n)$ と置く事に依り

$$F(x, t) = \sum_{n=0}^{\infty} P_n(t) \varepsilon(x-n) = \sum_{n \leq x} P_n(t)$$

となる。今 $P_n(t)$ は (26) で與へられるものとし此の様な一般の場合をも compound Poisson process と呼ぶ事にする。

$\Re(\zeta) \geq 0$ に対して

$$P(\zeta) = \int_0^{\infty} e^{-\zeta x} dU(x),$$

$$z = z(v) = \int_{-\infty}^{\infty} e^{ivx} dV(x)$$

と置けば、 $F(x, t)$ に對應する特性函数は

$$(40) \quad \begin{aligned} f(v, t) &= \int_{-\infty}^{\infty} e^{ivx} dF(x, t) = \int_{-\infty}^{\infty} e^{ivx} \sum_{n=0}^{\infty} P_n(t) dV_n(x) \\ &= \sum_{n=0}^{\infty} P_n(t) \int_{-\infty}^{\infty} e^{ivx} dV_n(x) = \sum_{n=0}^{\infty} z^n P_n(t) \\ &= \sum_{n=0}^{\infty} z_n \int_0^{\infty} \frac{(tx)^n}{n!} e^{-tx} dU(x) = \int_0^{\infty} e^{-tx} \cdot e^{itz} dU(x) \\ &= P(t(1-z)). \end{aligned}$$

$U(x), V(x)$ が有限の k 次の absolute moment を持つとすれば

$$(41) \quad P(\zeta) = 1 + \sum_{n=1}^k \frac{a_n}{n!} (-\zeta)^n + o(|\zeta|^k)$$

$$(42) \quad z = z(v) = 1 + \sum_{n=1}^k \frac{c_n}{n!} (iv)^n + o(v^k)$$

但し

$$a_n = \int_0^{\infty} x^n dU(x), \quad c_n = \int_{-\infty}^{\infty} x^n dV(x).$$

(40), (41), (42) から $f(v, t)$ は次の形をとる。

$$f(v, t) = 1 + \sum_{n=1}^k \frac{\nu_n}{n!} (iv)^n + o(v^k).$$

即ち $F(x, t)$ は有限な k 次の absolute moment を持ち ν_n は $F(x, t)$ の n 次の moment になる。

而して ν_n, a_n, c_n の間の關係は

$$f(v, t) = P(t(1-z))$$

から容易に

$$(43) \quad \nu_1 = ta_1c_1, \quad \nu_2 = ta_1c_2 + t^2a_2c_1^2, \dots$$

$F(x, t), U(x)$ の semi-invariant を λ_n, γ_n とすれば

$$\log f(x, t) = \log P(t(1-z))$$

から容易に次の關係が得られる。

$$(44) \quad \lambda_1 = t\gamma_1c_1, \quad \lambda_2 = t\gamma_1c_2 + t^2\gamma_2c_1^2, \dots$$

random change の大さ 1 なる彷徨過程では $c_n=1$ である。分布 $U(x)$ の平均値が有限なるとき

は Operational time scale $\lambda = \sum_{n=1}^{\infty} n P_n(t) = a_1 t$ を用ひる事にすれば $U(x)$ は平均値 1 なる分布 $U(a_1 x)$ に變換されるから、始めから $a_1 = 1$ として一般である。 $(x_i - tc_i)/\sqrt{tc_i}$ の平均値は 0, 標準偏差の自乗は $1 + r_2 tc_i/c_2$ である。而して後者が 1 になるのは $r_2 = 0$ 又は $c_1 = 0$ の場合に限る。而して特性函数は

$$(45) \quad e^{-\frac{c_1}{\sqrt{c_2}} iv \sqrt{t}} f\left(\frac{v}{\sqrt{c_2 t}}, t\right) = e^{-\frac{c_1}{\sqrt{c_2}} iv \sqrt{t}} \cdot P\left(t\left[1 - z\left(\frac{v}{\sqrt{c_2 t}}\right)\right]\right) \\ = e^{-\frac{c_1}{\sqrt{c_2}} iv \sqrt{t}} P\left(-\frac{c_1}{\sqrt{c_2}} iv \sqrt{t} + \frac{v^2}{2}\right) + o(1), (t \rightarrow \infty).$$

彷徨過程が齊時なるとき。而してその時に限り $r_2 = 0$, $U(x) = \varepsilon(x-1)$, $P(\zeta) = e^{-\zeta}$, 従て $t \rightarrow \infty$ なるとき (45) は v に就き一様に

$$P(v^2/2) = e^{-\frac{v^2}{2}}$$

に收斂する。又 $c_1 = 0$ なるときは (45) は

$$P\left(\frac{v^2}{2}\right) = \int_0^{\infty} e^{-\frac{v^2}{2}\xi} dU(\xi)$$

に收斂する。而して之は分布函数

$$\phi(x) = \int_0^{\infty} \phi\left(\frac{x}{\sqrt{\xi}}\right) dU(\xi)$$

の特性函数である。但し

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{y^2}{2}} dy,$$

故に $(x_i - tc_i)/\sqrt{tc_i}$ の limit distribution は齊時なるとき正規分布 $\phi(x)$ であり, $c_1 = 0$ なるとき $x_i/\sqrt{tc_i}$ の limit distribution は $\phi(x)$ である。前者は良く知られた事實である (Cramér [2] p. 97).

齊時彷徨過程に對しては x_i の分布函数の漸近展開に關する Cramér の結果が知られてゐる (Cramér [2], 定理 25, 30). 著者はこの Cramér の結果を用ひて (A) $c_1 = 0$ なる一般の compound Poisson process 及び (B) $c_1 \neq 0$ なる Pólya process に對して次の結果 (A), (B) を得てゐる。

(A) $k \geq 3$, $\varepsilon > 0$ に對して

$$U(x) = O(x^{\frac{1}{2}k-1+\varepsilon}), \quad \int_{-\infty}^{\infty} |x|^k dV(x) < +\infty, \\ \int_{-\infty}^{\infty} x dV(x) = 0, \quad \int_{-\infty}^{\infty} V(x) dx > 0$$

ならば

$$F(x\sqrt{c_2 t}, t) = \phi(x) + \sum_{n=1}^{k-1} \frac{\psi_{3n-1}(x)}{t^{n/2}} + Q_k(x, t),$$

但し $\psi_{3n-1}(x)$ は x の ν の多項式 $p_\nu(t)$ を以て

$$\psi_{3n-1}(x) = \int_0^{\infty} \xi^{-\frac{n}{2}} p_{3n-1}\left(\frac{x}{\sqrt{\xi}}\right) e^{-\frac{x^2}{2\xi}} dU(x)$$

で與へられ, $Q_k(x, t)$ は k, t に無關係な或る常數 M に對して

$$|Q_k(x, t)| < \frac{M}{t^{k/2-1}}$$

を満足する。

(B)

$$U'(x) = w_b(x) = \frac{1}{b} \frac{\left(\frac{x}{b}\right)^{1/b-1}}{\Gamma\left(\frac{1}{b}\right)} e^{-\frac{x}{b}}, \quad x \geq 0, \quad 0 < b < 1,$$

$$\int_{-\infty}^{\infty} |x|^3 dV(x) < \infty, \quad \int_{-\infty}^{\infty} x dV(x) = 1, \quad \int_{-\infty}^{\infty} V'(x) dx > 0$$

ならば

$$F(xt, t) = \int_0^x w_b(u) du + \frac{1}{2} \frac{c_2}{b} \frac{1}{t} (1-x) w_b(x) + R(x, t),$$

而して t, x に無関係な或常數 M に對して

$$|R(x, t)| < M \left(\frac{1}{t^2} + \frac{\log t}{t^{1/b}} \right).$$

§6. 應 用

應用の中主なものを掲げる. 次の Table A, B は疾病保險會社 *Eir* の記録に依るもので, Table A, B は夫々二つの Period:-(1) 保險年度 3-7, (2) 保險年度 8-12 及び (1) 曆年度 1931-33, (2) 曆年度 1934-36 に於けるもので

N = 契約者の總數,

$N_n(t)$ = $(0, t)$ 期間内に n 回保險事故 (罹病) を生じた契約者數,

$N_{m:n}(s, t)$ = 期間 $(0, s)$ 内に m 回保險事故を生じ, 期間 (s, t) 内に $n-m$ 回保險事故を生じた契約者の數,

とすれば, 今迄の記號との對應 (~印で示す) は次の如くである.

$$N_n = \sum_{m=0}^n N_{m,n}, \quad N = \sum_n N_n$$

$$\frac{N_n(t)}{N} \sim P_n(t), \quad \frac{N_{m,n}}{N_m} \sim P_{m,n}(s, t)$$

$$\frac{N_{m,n}}{N_n} \sim Q_{m,n}(s, t) = \frac{P_m(s) \cdot P_{m,n}(s, t)}{P_n(t)}.$$

而して Table A, B に於て $N_{m:n}$ は m 行 $n-m$ 列にある group の frequency とする. Table A, B が compound Poisson process からの sample であるかどうかを次の形に於て判定する.

(1) Parameter の分布函數 $U(x)$ は stable であるか, 即ち相異なる Period (1), (2) に對して時間に無関係な $U(x)$ を考へ得るか? (moment criterion).

(ii) inverse probability $Q_{m:n}(s, t)$ は二項分布と看做し得るか? (binomial criterion).

moment criterion. moment の estimate を求めるのに等式

$$\sum_{n=0}^{\infty} z^n P_n(t) = P(t(1-z))$$

の兩邊を z に就き k 回微分して得られる

$$(46) \quad t^k a_k = \sum_{n=0}^{\infty} n(n-1) \cdots (n-k+1) P_n(t)$$

を用ひる. 各 Period の長さを契約者の數の平均値とする. 從て (46) から分る様に a_1 の estimate を常に 1 とする. 而して問題を $U(x)$ の標準偏差の自乗の比較に局限する. $U(x)$ の標準偏差の自

Table A

Period 2

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	Total	Line means	
0	324	141	70	26	17	4	1	2								585	0.80	0
1	158	97	60	32	12	4	2		1	1			1	1		369	1.18	1
2	42	39	43	20	17	13	4	1								179	1.95	2
3	26	25	21	20	11	15	5	2	1							126	2.37	3
4	4	9	14	10	6	10	4	3	1	1	1					63	3.38	4
5	2	2	4	11	5	1	2	2	1		1			1		32	3.91	5
6	1	5	7	4	2	6	2	3			1					31	3.61	6
7	2	1		1	1	2	1			1		1				10	4.30	7
8				2	3		1	2	2		1					11	5.82	8
9	1	1		1			1			1						5		9
10				1	1		1				1					4		10
11																		11
12																		12
13											1					1	5.91	13
14																		14
15																		15
16																		16
17															1	1		17
Total	560	320	219	128	75	55	24	15	7	3	6	1	1	2	1	1417	1.53	
Column means	0.67	1.07	1.49	2.29	2.41	3.09	4.04	4.27	5.43	6.36						1.39		

Table B

Period 2

	0	1	2	3	4	5	6	7	8	9	10	Total	Line means	
0	187	94	39	12	7							339	0.70	0
1	91	110	71	26	11	2	1		1			313	1.27	1
2	51	53	36	18	4	5						167	1.32	2
3	28	32	23	24	7	2						116	1.62	3
4	6	14	11	9	7	1	1				1	50	2.24	4
5	2	8	11	6	4	1		1				33	2.30	5
6		5	3	2	4		1					15	2.60	6
7	1	3	4	2	1	2	1					14	2.64	7
8			2	1		1	2					6	3.56	8
9		1										1		9
10			1			1						2		10
Total	366	320	201	100	45	15	6	1	1		1	1056	1.27	
Column means	0.87	1.46	1.91	2.34	2.64	3.93	4.89				1.54			

乗は (43) から $b = (\nu_2 - \nu_1) / \nu_1^2 - 1$ に等しい。今二つの Period (1), (2) に對して標準偏差の自乗を比較する前に二重の random sampling をやつてゐる事に注意する。即ち第一に $U(x)$ の規定する母集團から N 人を取り出してゐる事、第二に各個人に對して或る時間内の罹病回数を考へてゐる事

である。然し Table A, B では同じ個人の集団を考へてゐる事から第一の要素を無視してよい。即ち此の N 人に対して (t 時間内に罹病する) 確率分布 $(t^k)/n! e^{-t^k}$ を持つ確率變數 x_i ($i = 1, 2, \dots, N$) を考へ、又確率變數

$$\bar{v}_k = \frac{1}{N} \sum_{i=1}^N x_i^k$$

を定義し、 $\mathbf{u} = \bar{v}_2 - \bar{v}_1$ 、 $\mathbf{v} = \bar{v}_1^2$ と置けば準偏差の自乗 \mathbf{b} に對應する確率變數は

$$\mathbf{b} = \frac{\mathbf{u}}{\mathbf{v}} - 1$$

である。而して互に共通部分のない Period (1) (t_1) (2) (t_2) には互に獨立な確立變數 $(\mathbf{u}_1, \mathbf{v}_1)$ 、 $(\mathbf{u}_2, \mathbf{v}_2)$ が對應する。 t_1, t_2 對する \mathbf{b} の estimate を夫々 b_1, b_2 とし $a = |b_1 - b_2|$ と置けば

$$(47) \quad \Pr. \left(\left| \frac{\mathbf{u}_1}{\mathbf{v}_1} - \frac{\mathbf{u}_2}{\mathbf{v}_2} \right| > a \right) = \Pr. (z > 0) + \Pr. (z' < 0),$$

但し

$$z = \mathbf{u}_1 \mathbf{v}_2 - \mathbf{u}_2 \mathbf{v}_1 - a \mathbf{v}_1 \mathbf{v}_2$$

$$z' = \mathbf{u}_1 \mathbf{v}_2 - \mathbf{u}_2 \mathbf{v}_1 + a \mathbf{v}_1 \mathbf{v}_2.$$

此處で z, z' の近似的分布として正規分布を假定すれば

$$(48) \quad \Pr. (z > 0) \sim 1 - \phi \left(\frac{0 - E(z)}{D(z)} \right) = \phi \left(\frac{E(z)}{D(z)} \right)$$

$$\Pr. (z' < 0) \sim \phi \left(\frac{0 - E(z')}{D(z')} \right) = \phi \left(-\frac{E(z')}{D(z')} \right),$$

但し E, D は夫々平均値及び標準偏差とする。 N は Table A, B に於て夫々 1417, 1056 で $\alpha(1/N)$ の項は省略してよい事が分る。その結果

$$(49) \quad E(z) = -E(z') \sim a(t_1 t_2)^2$$

$$D^2(z) = D^2(z') \sim 4 \frac{(t_1 t_2)^2}{N} \left\{ a_3 t_1 t_2 (t_1 + t_2) + \frac{1}{2} a_2 (t_1^2 + t_2^2) - a_2^2 t_1 t_2 (t_1 + t_2) + a^2 t_1 t_2 (t_1 + t_2) \right\},$$

此處に a_k の estimate は (46) に於て $t = t_1 + t_2$ と置いて求める。

Table A から

$$N = 1417, \quad t_1 = 1.3945, \quad t_2 = 1.5328, \quad t_1 + t_2 = 2.9273$$

$$b_1 = 1.0410, \quad b_2 = 0.9344, \quad a = b_1 - b_2 = 0.1066$$

$$a_1 = 1, \quad a_2 = 1.9416, \quad a_3 = 5.7689$$

(49) に依て

$$E(z) = -E(z') \sim 0.4870, \quad D(z) = D(z') \sim 0.4678,$$

$$\frac{E(z)}{D(z)} = -\frac{E(z')}{D(z')} = -\frac{4870}{4678} = -1.041.$$

(47), (48) に依り

$$\Pr. \left(\left| \frac{\mathbf{u}_1}{\mathbf{v}_1} - \frac{\mathbf{u}_2}{\mathbf{v}_2} \right| > 0.1066 \right) \sim 0.28.$$

Table B から

$$\begin{aligned}
 N &= 1056, & t_1 &= 1.5388, & t_2 &= 1.2671, & t_1 + t_2 &= 2.8059 \\
 b_1 &= 5.661, & b_2 &= 0.3035, & a &= b_1 - b_2 &= 0.2626 \\
 a_1 &= 1, & a_2 &= 1.4571, & a_3 &= 2.7443
 \end{aligned}$$

故に今 0.05 を有意 (significance) の level にとれば, Table B に對しては標準偏差の自乗の deviation は有意であるが, Table A に對しては有意でない.

binomial criterion. Table A, B より計算した値 $N_{m,n}/N_n$ と之に對應する

$$Q_{m,n} = \binom{n}{m} \left(\frac{s}{t}\right)^m \left(1 - \frac{s}{t}\right)^{n-m}$$

を比較するのに, Pearson の χ^2 検査法を用ひる. 先づ

$$\chi_n^2 = \sum_{m=0}^n \frac{(N_{m,n} - N_n \cdot Q_{m,n})^2}{N_n \cdot Q_{m,n}} \quad (n=1, 2, \dots)$$

を計算し, 之等の値及びそれに対応する自由度を χ^2 の和の法則に従て加へると

Table A から

$$\chi^2 = 42.3, \quad \text{自由度 } 27, \quad P = 2-5\%,$$

Table B から

$$\chi^2 = 33.1, \quad \text{自由度 } 20, \quad P = 2-5\%.$$

今有意の level を 0.05 にとれば, binomial criterion に依れば, $U(x)$ が存在すると云ふ假設は何れの data に對しても捨てなければならない.

引用文献 (抜萃)

- (1) S. Bernstein [1] Sur les fonctions absolument monotones. Acta Math. Vol. 52 pp. 1—66 (1929).
- (2) Cramér [1] Ein Grenzproblem in der Spieltheorie. Zeitschr. f. angew. Math. u. Mech. B. 13 pp. 76—79 (1933).
- (3) Cramér [2] Randon Variables and Probability Distributions. Cambridge (1937).
- (4) W. Feller [1] Zur Theorie der Stochastischen Prozesse. Math. Annalen 113 pp. 113—160 (1936).
- (5) W. Feller [2] Completely monotone functions and sequences. Duke Math. Journ. pp. 661—674 (1939).
- (6) M. Greenwood and Yule [1] An Inquiry into the Nature of Frequency Distributions Journ. Roy. Stat. Soc. Vol. 83, pp. 255—279 (1920).
- (7) A. Kolmogoroff [1] Über die analytischen Methoden in der Wahrscheinlichkeitsrechnung. Math. Annalen 104 pp. 415—458 (1931).