J. Neyman : On a new class of "contagious" distributions applicable in Entomology and Bacteriology, Ann. Math. Statistics, 10(1939) 35-57.

吉原, 友吉
東北帝國大學理學部

出版情報：統計数理研究. 1 (1), pp.102-104, 1941-10-15. 統計科学研究会

論 著 紹 介

J. Neyman

On a new class of "contagious" distributions applicable in Entomology and Bacteriology, Ann. Math. Statistics, 10(1939) 35–57.

會員 吉 原 友 吉 (東北帝大理學部)

實驗野場に於ける分布問題には Poisson の分布法則の適用し難い場合がある. Pólya はこの種の Contagious distribution に對してある分布法則を與へた ("Sur quelques points de la théorie des probabilités" Ann. de l'Institut Henri Poincaré I. 1931. 117–162). 著者は昆蟲の幼蟲の分布に就いて三つの新しい型の分布法則を求めた.

實驗野 F は單位面積の Plot: P が M 個集つてゐるとする. F 內の點 (ξ, η) に在る産卵塊のある一つにて孵化して, 實驗の時まで生存せる幼蟲數 n を確率變數とし, その確率密度を $p(n)$ とする. (ξ, η) を確率變數とすれば $p(\xi, \eta)=\frac{1}{M}$ はその確率密度を表はす. (ξ, η) で孵化し實驗の際, 點 (x, y) にゐた幼蟲を考へる. (x, y) の確率密度を $f(x-\xi, y-\eta)$ とする. F 上に産卵塊は N 個ありとし, 第 i 番目にて孵化し, 實驗時に P 內でその中の k_i 匹の幼蟲が見出さるるとすれば, P 內にて合計 $X=\sum_{i=1}^{N} k_i$ 丈幼蟲が見出される. 依つて X の確率密度を求むるには k_i のそれを求むればよい. (ξ, η) で孵化して後 P 內に見出さるる確率は

$$P(\xi, \eta)=\iint_P f(x-\xi, y-\eta)\, dx\, dy$$

なる故, (ξ, η) で n 匹孵化し, その中 k 匹丈 P 內で見出さるる確率は

$$P\{k \mid n, \xi, \eta\}=\binom{n}{k} P^k(\xi, \eta)(1-P(\xi, \eta))^{n-k} d\xi\, d\eta$$

故に某産卵塊で n 匹孵化せる中 P 內で k 匹見出さるる確率は

$$P\{k \mid n\}=\binom{n}{k} \frac{1}{M} \iint_F P^k(1-P)^{n-k} d\xi\, d\eta .$$

これに $p(n)$ を乘じ n に就て和を作れば k がある定められた値となる確率が求まる. $P(\xi, \eta)>0$ なる F 內の部分 (及びその面積) を A とすれば, A 內にて孵化せる幼蟲は實驗時までに P 內に移動してゐるから, $P\{k\}=\sum_{n \geq 0} p(n) P\{k \mid n\}$ は (ξ, η) で n 匹孵化してその中 k 匹丈 P 內に來る確率を表はす. k の特性函數は

$$\phi_k(t)=\sum_{k \geq 0} e^{-itk} P\{k\}=1-\frac{A}{M}+\frac{1}{M} \iint_A \sum_{n=0}(P e^{it}+1-P)^n p(n)\, d\xi\, d\eta$$

單位面積に平均 m ケの孵卵塊があるとすれば $N=Mm$. $k_1, k_2, \cdots, k_N$ は獨立と考へれば，その和 X の特性函數は $\phi_x=\prod_{k=1}^{N}\phi_k(t)$ となる．$M\to\infty$ と假定すれば

$$\phi(t) = \lim_{M\to\infty}\phi_x(t) = \exp\left\{-Am\left(1-\frac{1}{A}\right)\int_A\sum_{n\geqq 0}(Pe^{it}+1-P)^n\,d\xi\,d\eta\right\}$$

因つて

$$P\{X=X'\} = \frac{1}{2\pi}\int_{-\pi}^{\pi}\phi(t)\,e^{-itX'}\,dt.$$

各產卵塊につき實驗時まで殘存せる幼蟲の平均數を λ とし $p(n)=e^{-\lambda}\cdot\frac{\lambda^n}{n!}$ とすれば，上式より

$$\log\phi(t) = Am\sum_{n=1}^{\infty}\frac{\lambda^n(e^{it}-1)^n}{n!}P_n,\qquad 但\quad P_n = \frac{1}{A}\iint_A P^n\,d\xi\,d\eta.$$

今 $P(\xi,\eta)$ を確率變數 Z とし $0<z<1$ とするときは，$F(z)=P\{P(\xi,\eta)<z\}/A$ は Z の分布函數であり，$P_n=\int_0^1 z^n dF$ より $AP_1=1$. 故に

$$(*)\ \log\phi(t) = m\lambda(e^{it}-1) + Am\sum_{n=2}^{\infty}\frac{\lambda^n(e^{it}-1)^n}{n!}P_n.$$

而して $m\lambda(e^{it}-1)$ は $p(x)=e^{-m\lambda}(m\lambda)^x/x!$ の特性函數の對數となる．又 $AP_n\leqq 1$, 特に $AP_n\to 0$ $(n\leqq 0)$ なるときは t に關して一様に

$$\log\phi(t)\to m\lambda(e^{it}-1).$$

以上のことより [I] m,λ が一定なるとき，$F(z)$ が $AP_n\to 0$ なる如く定まれば，t に關して一様に $\phi(t)=m\lambda(e^{it}-1)$ となり，X は Poisson の分布法則に從ふ．[II] (*) なる型の特性函數を有する S 個の獨立なる確率變數 $X_1, X_2, \cdots X_S$ に對してはその和 $Y=\sum_{i=1}^{s}X_i$ も亦同じ分布をなすことがわかる．次に $P(\xi,\eta)$ の型により三つの分布型に分ける．

Type A : (2-parameter). $P(\xi,\eta)=A^{-1}((\xi,\eta)\varepsilon A)$ なるときは $m_1=Am$, $m_2=\lambda/A$, $z=m_1e^{-m_2}$ とすれば

$$\phi_1(t) = \exp\left\{-Am\left(1-\exp\frac{\lambda(e^{it}-1)}{A}\right)\right\} = e^{-m_1}\sum_{k=0}^{\infty}\frac{m_2^k}{k!}e^{ikt}\sum_{n=0}^{\infty}\frac{z^n}{n!}n_k$$

依つて

$$P\{X=k\} = e^{-m_2}\frac{m_2^k}{k!}\sum_{t=0}^{\infty}\frac{2^t}{t!}t^k.$$

尚 G. Beall は次の關係を與へた．

$$P\{X=n+1\} = \frac{m_1m_2}{n+1}e^{-m_2}\sum_{t=0}^{n}\frac{m_2^t}{t!}P\{X=n-t\}.$$

m_1, m_2 は moment を以て實驗より求められる．即ち $m_1=\mu_1'/m_2$, $m_2=(\mu_2-\mu_1')/\mu_1'$. (3-parameter). $P(\xi,\eta)$ が A 內で B_1 及び B_2 なる値を夫々確率 1/2 でとるときは，$m_1=Am$, $m_2=\lambda B_1$, $m_3=\lambda B_2$ とするとき

$$p_2(t) = \frac{Am}{2}\left(e^{\lambda(e^{it}-1)B_1}+e^{\lambda(e^{it}-1)B_2}-2\right).$$

又 $\log\phi(t)=\phi_2(t)$ とすれば $\psi(z)=\phi(t)=\sum_{k=0}^{\infty} z^k P\{X=k\}$ より $P\{X=k\}$ が求められる．又

$$P\{X=n+1\} = \frac{m_1}{2(n+1)}\sum_{k=0}^{n}\frac{m_2^{k+1}e^{-m_2}+m_3^{k+1}e^{-m_3}}{k!}P\{X=n-k\}$$

Type B： $0<z<1$ で $dF/dz=p_1(z)=\frac{1}{2}A(0<z<\frac{2}{A})$，他で $=0$ $(A>2)$ とすれば

$$\psi_3(z) = e^{-m_1}e^{m_1\frac{e^{m_2(z-1)}-1}{m_2(z-1)}} = \sum_{n=0}^{\infty} z^n P\{X=n\},$$

但し $m_1=Am,\ m_2=2\lambda/A.$

Type C： 同様に $p_2(z)=\frac{2}{9}A^2(\frac{3}{A}-z)$ $(0<z<\frac{3}{A})$，他で $=0$ とすれば

$$\psi_4(z) = e^{-m_1}\exp m_1\frac{e^{m_2(z-1)}-1-m_2(z-1)}{m_2^2(z-1)^2}\sum_{n=0}^{\infty} z^n P\{X=n\}.$$

となる．

著者はここに求めたる式を用ひたるに次の如き結果を得た．

第1表： 木喰蟲の分布 (Distribution of European corn borers in 120 groups of 8 hills each)

第2表： 酵母菌の分布 (Distribution of yeast cells in 400 sq. of haemacytometer observed by "Student" (1907), Biometrika 5. 1907. (351-364)),

(尚 J. Neyman "Lectures and Conferences on Mathematical Statistics", 1938. 参照).

第 1 表

木喰蟲の數	頻度 Poisson の式	實測値	Type A
0	5.0	24	22.6
1	16.0	16	16.7
2	25.3	16	18.3
3	26.7	18	16.4
4	21.1	15	13.4
5	13.4	9	10.3
6	7.1	6	7.5
7	3.2	5	5.2
8	1.3	3	3.5
9	0.4	4	2.3
10	0.1	3	1.5
11	—	0	—
12	—	1	—
(以下)	—	—	2.3
m_1	—	—	2.178
m_2	—	—	1.454
P_{χ}^2			0.95

第 2 表

酵母菌の數	頻度 Poisson の式	實測値	Type A
0	202	213	214.8
1	138	128	121.3
2	47	37	45.7
3	11	18	13.7
4	—	3	3.6
5	—	1	0.8
(以下)	2	—	0.1
m_1	—	—	3.605
m_2	—	—	0.189
P_{χ}^2	—	>0.2	>0.1

Alf Guldberg

Les fonctions de fréquence discontinues et les séries statistiques,

Annales de L'innstitut Henri Poincaré, 3(1932), 229-278.

會員 丸 山 儀 四 郎 (九州帝大理學部)

此の論文の後半に於て，次の諸分布：

(1) Poisson の分布 $f(x) = \dfrac{\lambda^x e^{-\lambda}}{x!}$ $(x = 0, 1, 2, \cdots\cdots)$

(2) 二 項 分 布 $f(x) = {}_kC_x p^x q^{k-x}$, $p + q = 1$, $(x = 0, 1, 2, \cdots\cdots k)$

(3) Pascal の分布 $f(x) = {}_{k+x}C_k p^k q_x$, $p + q = 1$, $(x = 0, 1, 2, \cdots\cdots)$

を使用せんとする場合，その何れを用ふべきかに對する一つの判定法を與へてゐる．方法は (1), (2), (3) 何れに對しても同樣であるから，以下 (2) に就いて考へる．(2) は次の定差方程式を滿足する

(4)
$$f(x+1) = \frac{p}{1-p}\,\frac{k-x}{x+1}\,f(x),$$

或は

$$(1-p)(x+1)f(x+1) = pkf(x) - pxf(x).$$

兩邊に $(x+1)^n$ を掛けて右邊を x の羃に展開すると

$$(1-p)(x+1)^{n+1}f(x+1) = pkx^n f(x) + {}_nC_1 pkx^{n-1}f(x) + \cdots\cdots + pkf(x)$$
$$- px^{n+1}f(x) - {}^nC_1 px^n f(x) - \cdots\cdots - pxf(x).$$

σ_n を原點の周りの n 次の moment，即ち

$$\sigma_n = \sum_{x=0}^{\infty} x^n f(x)$$

とすれば，上の式の兩邊を x に就き 0 から k 迄加へる事に依り次の回歸式が得られる．

$$(1-p)\sigma_{n+1} = pk\sigma_n + {}_nC_1 pk\sigma_{n-1} + \cdots\cdots + pk$$
$$- p\sigma_{n+1} - {}_nC_1 p\sigma_n - \cdots\cdots - p\sigma_1.$$

特に $n = 0, 1$ と置き

$$\sigma_1 = kp, \quad \sigma_2 = pk\sigma_1 + pk - p\sigma_1 = p^2k^2 - p^2k + pk.$$

最後に求める式の形を簡單にするために semiinvariants，即ち

$$e^{\mu_1 it + \frac{\mu_2}{1\cdot 2}(it)^2 + \cdots\cdots} = \sum_{x=0}^{\infty} e^{ixt} f(x)$$

なる如き μ_1, μ_2 を用ふれば良く知られてゐる公式に依つて

$$\mu_1 = \sigma_1 = kp, \quad \mu_2 = \sigma_2 - \sigma_1^2 = kp(1-p), \quad \mu_2 < \mu_1.$$

故に

$$p = \frac{\mu_1 - \mu_2}{\mu_1}, \qquad k = \frac{\mu_1^2}{\mu_1 - \mu_2}.$$

之を (4) に代入すれば

$$\frac{f(x+1)}{f(x)}(x+1) + \frac{\mu_1 - \mu_2}{\mu_2} x = \frac{\mu_1^2}{\mu_2}.$$

(1) 及 (3) に對しても全く同じ定差方程式が得られる．但し (1) に於ては，$\mu_1 = \mu_2 = \lambda$，從つて

$$\frac{f(x+1)}{f(x)}(x+1) = \lambda$$

となる．又 (3) に對しては

$$\mu_1 = q\frac{k+1}{1-q}, \qquad \mu_2 = q\frac{k+1}{(1-q)^2}.$$

從つて $\mu_2 > \mu_1$ である．結局 (1), (2), (3) の各々に對して

$$a(x) = \left\{\frac{f(x+1)}{f(x)}(x+1) + \frac{\mu_1 - \mu_2}{\mu_2} x\right\} \Big/ \frac{\mu_1^2}{\mu_2} = 1.$$

而して (1) なるか，(2) なるか，(2) なるかに從ひ夫々

$$\mu_1 = \mu_2 = \lambda, \quad \mu_1 > \mu_2, \quad \mu_1 < \mu_2$$

である．故に (1), (2), (3) を用ひんとすれば，與へられた sample に就いて $a(x)$ の値を計算し，之が大略 1 に等しくなるか否かを見る．而して $a(x)$ の値が總べての x に對して大略 1 に等しい場合に μ_1, μ_2 が特に兩者の大小關係の何れを滿足してゐると見るべきかを知りて (1), (2), (3) の何れを用ふべきかを決定する．

例 1. Rutherford と Geiger は或る一定時間內に Polonium から出る α 粒子の數を觀測して次の結果を得た．x は α 粒子の數，$H(x)$ は x なる觀測値を得た觀測回數とする．

x	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
$H(x)$	57	203	383	525	532	408	273	139	45	27	10	4	0	1	1

之から計算すると，$\mu_1 = 3.88 = \lambda$ を得る．

$$a(x) = \frac{H(x+1)}{H(x)}(x+1) \Big/ 3.88$$

の $x = 0, 1, \cdots\cdots, 13$ に對する値を計算すると

$a(0) = 0.92$ $a(1) = 0.97$ $a(2) = 1.05$ $a(3) = 1.05$ $a(4) = 0.94$
$a(5) = 1.04$ $a(6) = 0.92$ $a(7) = 0.67$ $a(8) = 1.40$ $a(9) = 0.96$
$a(10) = 1.14$ $a(11) = 0$ $a(12) = \infty$ $a(13) = 3.62$.

依つて $\lambda=3.88$ として Poisson の分布 $f(x)=\dfrac{e^{-\lambda}\lambda^x}{x!}$ を用ふれば

x	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
$2608f(x)$	53	205	400	520	507	396	257	143	70	30	12	4	1	0	0

例 2. 2000 頭の牝豚の Müller の腺の數に就いて次の様であつた.

x	0	1	2	3	4	5	6	7	8	9	10
$H(x)$	15	209	365	482	414	277	134	72	22	8	2

之に對しては $\mu_1=3.5$, $\mu_2=2.8$, $\mu_1>\mu_2$ と見るべきである. 而して

$\alpha(0)=4.00$ $\alpha(1)=0.90$ $\alpha(2)=1.10$ $\alpha(3)=0.97$ $\alpha(4)=0.93$

$\alpha(5)=0.86$ $\alpha(6)=1.10$ $\alpha(7)=0.77$ $\alpha(8)=0.97$ $\alpha(9)=0.70$.

依つて (2) を用ふれば

x	0	1	2	3	4	5	6	7	8	9	10
200 (x)	41	177	364	468	423	286	150	62	21	6	1

例 3. v. Bortkiewicz に依ると 1869–1893 に於ける普魯西の子供の自殺者の數に就いて

x	0	1	2	3	4	5	6
$H(x)$	4	8	5	3	4	0	1

である, 但し x は某一年間に於ける自殺者數, $H(x)$ は自殺者 x を出した年の數. 之に就いては $\mu_1=1.96$, $\mu_2=2.46$, $\mu_2>\mu_1$. 而して

$\alpha(0)=1.3$ $\alpha(1)=0.7$ $\alpha(2)=0.9$ $\alpha(3)=0.3$ $\alpha(4)=0.1$

$\alpha(5)=\infty$ $\alpha(6)=0.8$.

依つて Pascal の分布を用ひると

x	0	1	2	3	4	5	6
$25f(x)$	4.0	6.2	5.5	3.6	2.0	0.9	0.4

v. Bortkiewicz が Poisson の分布で計算した結果に依ると

x	0	1	2	3	4	5	6
$25f(x)$	3.4	6.8	6.8	4.5	2.2	0.9	0.3

Pascal の分布に依る方が良好である.