二項分布函數，ポアソンの法則及びポリアーエッゲンベルガーの法則について

佐藤，良一郎
東京高等師範學校

出版情報：統計数理研究. 1 (1), pp.60-65, 1941-10-15. Research Association of Statistical Sciences

二項分布函數，ポアソンの法則及びポリアーエッゲンベルガーの法則について

會員 佐藤良一郎（東京高師）

1. 序言

よく知られてゐるやうに，0, 1, 2, ……, n, …… 中の何れかの値を取る變數 X があつて，負ならざる整數値 x を取る確率が

$$(1)\qquad P(x)=\frac{n!}{x!(n-x)!}p^x q^{n-x}$$

但し n は整數，$p>0$, $p+q=1$ あつて，$x>n$ に對しては $P(x)=0$ とする

で與へられるときには，X の分布法則或は確率法則は二項分布函數で表されるといひ，

$$(2)\qquad P(x)=\frac{e^{-\lambda}\lambda^x}{x!}$$

（但し λ は正の常數）

で與へられるときには，ポアソン (Poisson) の法則で表される，

$$(3)\qquad P(x)=\frac{h(h+d)\cdots(h+\overline{x-1}d)}{x!}(1+d)^{-\frac{h}{d}-x}$$

但し h, d は何れも正の常數で $P(0)=(1+d)^{-\frac{h}{d}}$ とする

で與へられるときには，ポリアーエッゲンベルガー (Polya-Eggenberger) の法則で表されるといふ．

これ等三個の分布函數は次のやうに考へると，何れも一つの分布函數の特別な場合と見られる．

2. 擴張されたポリアーエツゲンベルガーの法則

次の式

$$(4)\qquad \left(1-\frac{d}{1+d}\right)^{-\frac{h}{d}}$$

（但し h, d は常數である）

が，二項定理による展開 $\left(-\dfrac{d}{1+d}\right.$ に關しての$\left.\right)$ 可能なるときは，その一般項は

$$(5)\qquad \phi(r)=\frac{h(h+d)\cdots(h+\overline{r-1}d)}{x!(1+d)^r}$$

で表されるから，h, d を如何なる常數としようとも

$$(6)\qquad \sum_{x=0}^{\infty}\phi(x)(1+d)^{-\frac{h}{d}}\equiv\sum_{x=0}^{\infty}\frac{h(h+d)\cdots(h+\overline{x-1}d)}{x!}(1+d)^{-\frac{h}{d}-x}$$

が收斂するならば，その値は1となるべきである．實際このときには，

$$\sum_{x=0}^{\infty}\phi(x)(1+d)^{-\frac{h}{d}} = (1+d)^{-\frac{h}{d}}\sum_{x=0}^{\infty}\frac{h(h+d)\cdots(h+\overline{x-1}d)}{x!\,(1+d)^x} \tag{7}$$

$$= (1+d)^{-\frac{h}{d}}\left(1-\frac{d}{1+d}\right)^{-\frac{h}{d}} = 1.$$

但し便利のために，$-\dfrac{h}{d}=n$（正整數）なるときは，$x>n$ なる如き整數値に對しては $\phi(x)=0$ であり，又 $\phi(0)=1$ であるとする．

そこで，x の函數

$$P(x) = \phi(x)(1+d)^{-\frac{h}{d}} = \frac{h(h+d)\cdots(h+\overline{x-1}d)}{x!}(1+d)^{-\frac{h}{d}-x} \tag{8}$$

が，上掲の條件を滿足さしてゐる上に $x=0, 1, 2, \cdots\cdots$ に對して決して負にならないときには，これを擴張されたポリアーエッゲンベルガーの法則と假稱することとする．そして，變數 X が負ならざる整數値 x を取る確率が (8) で與へられるとき，X の分布法則或は確率法則は擴張されたポリアーエッゲンベルガーの法則（略して擴張された $P.E.$ 法則）で表されるといはう．

以上のやうに考へると，二項分布函數は擴張された $P.E.$ 法則の特別の場合となる．實際 $-h/d=n$（正整數）とするときは，

$$P(x) = \phi(x)(1+d)^{-\frac{h}{d}} = \frac{-nd(-nd+d)\cdots(nd+\overline{x-1}d)}{x!\,(1+d)^x}(1+d)^n$$

$$= \frac{n(n-1)\cdots(n-x+1)}{x!}(-d)^x(1+d)^{n-x}.$$

$P(x)$ は決して負とならないのであるから，$0<-d=p<1$ であるより外ないことは明かである．從つて

$$P(x) = \frac{n(n-1)\cdots(n-x+1)}{x!}p^x(1-p)^{n-x} \qquad (0<p<1)$$

即ち二項分布函數である．

ポアソンの法則は，擴張された $P.E.$ 法則に於て $d\to 0$ ならしめたときの極限として得られることは容易に知られる．

尙ほ $h/d=n$ なるときは

$$P(x) = \frac{n(n+1)\cdots(n+x-1)}{x!}p^x(1-p)^n \qquad \left(0<p=\frac{d}{1+d}<1\right)$$

を得る．これは形式の上からはパスカルの法則に似てゐるが，パスカルの法則に於ては上式の n が $n-1$ である．

3. 擴張された P.E. 法則の最初の二つのモーメント

擴張された $P.E.$ 法則の第一次・第二次のモーメントをそれぞれ μ_1, μ_2 とすると，次のやうに計算される．

I] $$\mu_1 = \sum_x x \cdot \frac{h(h+d)\cdots(h+\overline{x-1}d)}{x!\,(1+d)^x}(1+d)^{-\frac{h}{d}}$$

$$= h\sum_{x'} \frac{(h+d)(h+2d)\cdots(h+x'd)}{x'!\,(1+d)^{x'}}(1+d)^{-\frac{h+d}{d}} \qquad (x'=x-1)$$

(9) $$= h,$$

II] $$\mu_2 = \sum_x x^2 \cdot \frac{h(h+d)\cdots(h+\overline{x-1}d)}{x!\,(1+d)^x}(1+d)^{-\frac{h}{d}}$$

$$= \sum_x \Big\{x(x-1)+x\Big\}\frac{h(h+d)\cdots(h+\overline{x-1}d)}{x!\,(1+d)^x}(1+d)^{-\frac{h}{d}}$$

$$= h(h+d)\sum_{x'} \frac{(h+2d)(h+3d)\cdots(h+2d+x'd)}{x'!\,(1+d)^{x'}}(1+d)^{-\frac{h+2d}{d}} + h$$

(10) $$= h(h+d) + h \qquad (x'=x-2).$$

この結果から

$$\sigma^2 = \sum_x (x-h)^2 \frac{h(h+d)\cdots(h+\overline{x-1}d)}{x!\,(1+d)^x}(1+d)^{-\frac{h}{d}}$$

(11) $$= h(h+d) + (1-2h)h + h^2 = h(1+d)$$

なることが容易にわかる．従つて

(12) $$d = \frac{\sigma^2}{h} - 1$$

これは，ポリアーエッゲンベルガーの法則の場合に於ては傳染常數 (Contagion-Parameter) と呼ばれてゐるところのものである．

4. 擴張された P.E. 法則に對應する鑑別規準

觀察に依つて得た度數分布に，擴張された P.E. 法則が當嵌るかどうかを鑑別する規準としては次の等式 (13) を用ひることが出來る．

(8) から

$$\frac{P(x+1)}{P(x)} = \frac{h+xd}{(x+1)(1+d)}$$

従つて

$$\frac{P(x+1)}{P(x)}(x+1) = \frac{d}{1+d}x + \frac{h}{1+d}$$

を得るが (12) により $d/(1+d) = (\sigma^2-h)/\sigma^2$, $h/(1+d) = h^2/\sigma^2$ であるから

$$\frac{P(x+1)}{P(x)}(1+x) = \left(1-\frac{h}{\sigma^2}\right)x + \frac{h^2}{\sigma^2}$$

この鑑別規準を用ひるには次のやうにすればよい．即ち實際觀察に依つて

x	0	1	2	3 ………………	k	計
$f(x)$	$f(0)$	$f(1)$	$f(2)$	$f(3)$ ………………	$f(k)$	N

（$f(x)$ は x に對應する度數を表す）

なる度數分布を得たとするならば，先づ

$$\bar{x}=\frac{1}{N}\sum x\ ,\qquad s^2=\frac{1}{N}\sum(x-\bar{x})^2$$

を求め，$x=0,1,2,\cdots\cdots,k$ に對して

$$\frac{f(x+1)}{f(x)}(1+x)\qquad と\qquad \left(1-\frac{\bar{x}}{s^2}\right)x+\frac{h^2}{s^2}$$

が略ぼ相等しいかどうかを見定める．若し相等しと見られないときには，吾々の度數分布には，擴張された P.E. 法則は當嵌まらないのである．略ぼ相等しいと見得べきときには，吾々の度數分布には，先づ擴張された P.E. 法則が當嵌まるものと見られる．そこで $\bar{x}=h$, $\frac{s^2}{x}-1=d$ と置いて

$$F(x)=N\frac{h(h+d)\cdots(h+\overline{x-1}d)}{x!}(1+d)^{-\frac{h}{d}-x}\tag{15}$$

を作り，$x=0,1,2,\cdots\cdots,k$ に對する $F(x)$ の値を算出するならば，これが所謂理論的の度數或は期待度數となる．尤もこれが實際によく當嵌つてゐるか（確率論的に）どうかを檢べるには，別に備はるところの判定法，例へば χ^2 を用ひる檢査法の如きものに依らなければならない．

次に實例を示さう．

例 1. 中學校第一學年の生徒に，簡單な數計算の問題12個を課し，各生徒につき誤つて答へた問題の數（誤答數といはう）を調べて次のやうな結果を得た．

誤答數 (x)	0	1	2	3	4	5	6	7	8	9	10	11	12
生徒數 ($f(x)$)	267	517	532	484	299	217	105	70	25	14	5	—	—

$\bar{x}=2.677,\quad s^2=3.706,\quad N=2544.$

これに前述の鑑別規準を適用してみると次のやうになる．

		0	1	2	3	4	5	6	7	8	9	10	11	12
(1)	x	0	1	2	3	4	5	6	7	8	9	10	11	12
(2)	$\frac{f(x+1)}{f(x)}(x+1)$	1.873	2.058	2.729	2.451	3.629	2.904	4.669	2.856	5.040	3.571	—	—	—
(3)	$\left(1-\frac{\bar{x}}{s^2}\right)x+\frac{\bar{x}^2}{s^2}$	1.933	2.211	2.489	2.767	3.045	3.323	3.601	3.879	3.157	4.435	—	—	—

(1) 欄の數字と (2) 欄の數とを比べてみると，大體相等しいと見られるので，$d=\frac{s^2}{\bar{x}}-1=0.384$ を算出し，(15) に於ける N, h, d にそれぞれ 2544, 2.677, 0.384 を代入して，$x=0,1,2,\cdots\cdots,k$

に對應する $F(x)$ の値を算出してみると，次のやうになる．

x	0	1	2	3	4	5	6	7	8	9	10	11	12	計
$f(x)$	276	517	532	484	299	217	105	70	25	14	5	—	—	2544
$F(x)$	264.5	516.4	569.4	473.2	328.2	199.4	110.4	56.7	27.5	12.7	5.6	2.3	1.0	2567.9
												9.9		

$\chi^2 = 12.862,\quad f = 10,\quad P > 0.20$

χ^2 を用ひる検査法に依つて検べてみると，上表の欄外に記入してゐるやうな値となるから，先づ良く當嵌つてゐると見られる．從つて一般に，中學校第一學年の生徒の計算力といふものを誤答數の側から觀察すれば，ポリアーエッゲンベルガーの法則に從つて分布するものと推定してもよからう（上の例に於ては h, d 共に正である）．

__注意.__ 上掲の度數分布には，二項分布函數もポアソンの法則も共に當嵌り方が良くないのである．といふのは，二項分布函數が良く當嵌まるならば

$$d = \frac{\sigma^2}{h} - 1 = \frac{npq}{np} - 1 = q - 1 < 0 \qquad (0 < q < 1)$$

でなければならないし，ポアソンの法則が當嵌るならば，

$$d = 0$$

でなければならないが，上掲の度數分布に於ては $d = 0.384$ であつて，負でなく又零にも遠いからである．

__例 2.__ 骰子 12 個を 412 回投げて毎回六ノ目を出した骰子の數 x を觀察して，次の度數分布表を得た．

x	0	1	2	3	4	5	6	7	8	9	10	11	12	計
$f(x)$	45	115	118	80	38	12	2	1	1	—	—	—	—	412

この度數分布に對しては，問題にいふ事象の性質上，二項分布を當嵌めるのが自然であるが，毎回12個の骰子を投げたといふ事實には拘泥せず，唯與へられた一つの度數分布であるとして，上述の擴張された P.E. 法則を當嵌めるつもりで進行すると，次のやうな結果を得る．

$$\bar{x} = 2.015\,, \qquad s^2 = 1.767\,.$$

$\bar{x}, s^2$ をそれぞれ h, σ^2 に等しいと置くと

$$h = 2.015\,, \qquad \sigma^2 = 1.767\,, \qquad d = \frac{\sigma^2}{h} - 1 = -0.123$$

$$\frac{\sigma^2 - h}{\sigma^2} = -0.140\,, \qquad \frac{h^2}{\sigma^2} = 2.297$$

となる．

そこで，$f(x+1)(1+x)/f(x)$ 及び $\left(1 - \frac{h}{\sigma^2}\right)x + \frac{h^2}{\sigma^2}$ の $x = 0, 1, 2, \cdots\cdots$ に對する値を算出して比較してみると，次のやうになつて，主要部分に於て兩者は略ぼ相等しいと見られる

x	0	1	2	3	4	5	6	7	8
$\frac{f(x+1)}{f(x)}(x+1)$	2.556	2.052	2.034	1.900	1.579	1.000	3.500	8.000	—
$\left(1-\frac{h}{\sigma^2}\right)x+\frac{h^2}{\sigma^2}$	2.297	2.157	2.017	1.876	1.736	1.596	1.456	1.315	—

それで $N=412,\ h=2.015,\ d=-0.123$ と置いて，$x=0, 1, 2, \cdots\cdots$ に對應する

$$F(x) = \frac{N\cdot h(h+d)\cdots(h+\overline{x-1}d)}{x!}\,(1+d)^{-\frac{h}{d}-x}$$

の値を算出して，觀察の結果と對照すると次表の通りである．

x	0	1	2	3	4	5	6	7	8	9	10	11	12	計
$f(x)$	45	115	118	80	38	12	2	1	1	—	—	—	—	412
							14							
$F(x)$	47.99	110.24	118.88	79.92	37.44	13.00	3.46	0.72	0.12	0.01	—	—	—	411.78
								17.23						

$\chi^2=1.012,\ f=5,\ P>0.95$

χ^2 檢査法によつて當嵌まりの良さを檢べてみると，上表欄外に記入したやうに，$\chi^2=1.012$ でこれに對する P の値即ち χ^2 が 1.012 を超える確率は 0.95 より大きい．故に良く當嵌つてゐるといふべきである．

因に，二項分布函數を用ひ，x に對する度數 $F(x)$ を

$$F(x) = N\frac{n(n-1)\cdots(n-x+1)}{x!}\,q^x p^{n-x}$$

$$= N\cdot\frac{n(n-1)\cdots(n-x+1)}{x!}\left(\frac{h}{n}\right)^x\left(\frac{n-h}{n}\right)^{n-x}$$

で表はし，これに於ても $N=412,\ n=12,\ h=2.15$ と置いて，$x=0, 1, 2, \cdots\cdots$ に對應する値を算出して上の結果と對照すると次のやうになる．$f(x)$ は實際觀察によるもの，$F_1(x)$ は擴張された P.E. 法則，$F_2(x)$ は二項分布函數によるものである．

x	0	1	2	3	4	5	6	7	8	9	10	11	12	計
$f(x)$	45	115	118	80	38	12	2	1	1	—	—	—	—	412
$F_1(x)$	47.99	110.24	118.88	79.92	37.44	13.00	3.46	0.72	0.12	0.01	—	—	—	411.98
$F_2(x)$	45.43	109.90	122.20	81.97	37.29	12.32	2.84	0.49	0.06	—	—	—	—	412.10

上の表を見てもわかるやうに，$F_1(x)$ の系列と $F_2(x)$ の系列との間には，原系列に對する當嵌りの良さには甲乙がない．