今野　(2011 年 11 月 08 日)

情報統計学の問題 (その 2) の解答例

$\boxed{\text{問題 1}}$ (問 4.13) $\{X_n\}_{n=1}^{\infty}$ を確率変数列とし，各 X_n は確率関数

$$f_{X_n}(x) = \begin{cases} \frac{1}{n} & (x = c+n), \\ 1 - \frac{1}{n} & (x = c), \\ 0 & (\text{その他}) \end{cases}$$

を持つとする．ただし，c は定数とする．このとき，以下の問いに答えよ．

(1) $X_n \xrightarrow{P} c$ を示せ．

(2) 期待値の定義に従い，$\lim_{n\to\infty} \mathbb{E}[X_n]$ を求め，$\mathbb{E}[X_n] \not\to c$ であることを確認せよ．

解答. (1) 任意の正の数 $\epsilon > 0$ を取る．$1/n < \epsilon$ かつ $n > \epsilon$ なる大きな n に対して，X_n の確率関数の形に注意すれば，$\{X_n > c + \epsilon\}$ となる事象のなで，正の確率をもつ事象は $\{X_n = c + n\}$ しかないので，

$$\mathbb{P}(X_n > c + \epsilon) = \mathbb{P}(X_n = c + n) = \frac{1}{n}; \qquad \mathbb{P}(X_n < c - \epsilon) = 0$$

$$\begin{aligned} \mathbb{P}(|X_n - c| > \epsilon) &= \mathbb{P}(\{X_n - c > \epsilon\} \cup \{X_n - c < -\epsilon\}) \\ &= \mathbb{P}(X_n - c > \epsilon) + \mathbb{P}(X_n - c < -\epsilon) \qquad (\text{ふたつの事象は排反だから}) \\ &= \mathbb{P}(X_n > c + \epsilon) + \mathbb{P}(X_n < c - \epsilon) = \frac{1}{n} \end{aligned}$$

したがって，

$$\lim_{n\to\infty} \mathbb{P}(|X_n - c| > \epsilon) = 0$$

から $X_n \xrightarrow{P} c$ を示せた．

(2) 離散型確率変数の期待値の定義から

$$\begin{aligned} \mathbb{E}[X_n] &= (c+n)\mathbb{P}(X_n = c+n) + c\mathbb{P}(X_n = c) \\ &= (c+n) \times \frac{1}{n} + c \times \left(1 - \frac{1}{n}\right) = c + 1 \end{aligned}$$

なので，

$$\lim_{n\to\infty} \mathbb{E}[X_n] \neq c$$

がわかる．

$\boxed{\text{問題 2}}$ (問 4.13) 確率変数列 $\{X_n\}_{n=1}^{\infty}$ において，各 X_n(問題はタイプミス) は母数 n のポアソン分布に従うとする．すなわち，

$$\mathbb{P}(X_n = k) = \frac{n^k}{k!} e^{-n}, \qquad K = 0, 1, \ldots$$

である．

(1) $\mathbb{E}[X_1] = 1$ を示せ．ただし，$e = \sum_{k=0}^{\infty} \frac{1}{k!}$ を用いてよい．

(2) $\mathbb{E}[X_1(X_1 - 1)] = 1$ を示せ．

(3) $\mathbb{VAR}[X_1] = \mathbb{E}[X_1^2] - \{\mathbb{E}[X_1]\}^2 = 1$ を示せ．

(4) Z_1 と Z_2 が独立に母数 1 のポアソン分布に従うとき，$Z_1 + Z_2$ は母数 2 のポアソン分布に従うことを示せ．

ヒント：$\mathbb{P}(Z_1 + Z_2 = k) = \sum_{\ell=0}^{k} \mathbb{P}(Z_1 = \ell,\, Z_2 = k - \ell) = \sum_{\ell=0}^{k} \mathbb{P}(Z_1 = \ell)\mathbb{P}(Z_2 = k - \ell)$ となることと二項定理 $2^k = \sum_{\ell=0}^{k} {}_kC_\ell$ を用いる．

(5) $X_n/n \xrightarrow{P} 1$ を大数の法則を用いて示せ．ヒント（このヒントは (6) 用である）：$Z_1, Z_2, \ldots, Z_n$ が独立に母数 1 のポアソン分布に従うとき，上の問いの結果から $Z_1 + Z_2 + \cdots + Z_n$ が母数 n のポアソン分布に従うことを利用する．

(6) 中心極限定理を用いて，

$$\frac{X_n - n}{\sqrt{n}} \xrightarrow{d} N(0,\, 1)$$

を示せ．

解答. (1) 期待値の定義から

$$\mathbb{E}[X_1] = \sum_{k=0}^{\infty} \mathbb{P}(X_1 = k) = e^{-1} \sum_{k=0}^{\infty} k \times \frac{1}{k!} = e^{-1} \sum_{k=1}^{\infty} \frac{1}{(k-1)!} = e^{-1} \sum_{k=0}^{\infty} \frac{1}{k!} = e^{-1} \times e = 1$$

(2) 期待値の定義から

$$\begin{aligned}\mathbb{E}[X_1(X_1 - 1)] &= \sum_{k=0}^{\infty} k(k-1) \times \mathbb{P}(X_1 = k) = e^{-1} \sum_{k=0}^{\infty} k(k-1) \times \frac{1}{k!} = e^{-1} \sum_{k=2}^{\infty} \frac{1}{(k-2)!} \\ &= e^{-1} \sum_{k=0}^{\infty} \frac{1}{k!} = e^{-1} \times e = 1\end{aligned}$$

(3) 分散公式と期待値の性質より

$$\mathbb{VAR}[X_1] = \mathbb{E}[X_1^2] - \{\mathbb{E}[X_1]\}^2 = \mathbb{E}[X_1(X_1 - 1)] + \mathbb{E}[X_1] - \{\mathbb{E}[X_1]\}^2 = 1$$

(4)

$$\begin{aligned}
\mathbb{P}(Z_1+Z_2=k) &= \sum_{\ell=0}^{k}\mathbb{P}(Z_1=\ell,\, Z_2=k-\ell) \\
&= \sum_{\ell=0}^{k}\mathbb{P}(Z_1=\ell)\mathbb{P}(Z_2=k-\ell) \qquad \text{（Z_1 と Z_2 の独立性より）} \\
&= \sum_{\ell=0}^{k}\frac{1}{\ell!}e^{-1}\frac{1}{(k-\ell)!}e^{-1} \\
&= \frac{1}{k!}e^{-2}\sum_{\ell=0}^{k}\frac{k!}{\ell!(k-\ell)!} \\
&= \frac{1}{k!}e^{-2}(1+1)^k \qquad \text{（二項定理 $(a+b)^k=\sum_{\ell=0}^{k}a^{\ell}b^{k-\ell}$ から）} \\
&= \frac{2^k}{k!}e^{-2}
\end{aligned}$$

となるので，Z_1+Z_2 は母数 2 のポアソン分布に従うことがわかる．
(5) $Z_1, Z_2, \ldots, Z_n$ は独立同一に母数 1 のポアソン分布に従い，Z_1 は非負値確率変数であることに注意すれば，$\mathbb{E}[|Z_1|]=\mathbb{E}[Z_1]=1$ となり，$\mathbb{E}[|Z_1|]$ は有限なので，大数の法則を用いることができるので，

$$\frac{1}{n}(Z_1+\cdots+Z_n)\xrightarrow{P}\mathbb{E}[Z_1]=1$$

がわかる．
(6) $Z_1+\cdots Z_{n-1}$ を母数 $n-1$ のポアソン分布になることを仮定すると

$$\begin{aligned}
\mathbb{P}(Z_1+\cdots+Z_{n-1}+Z_n=k) &= \sum_{\ell=0}^{k}\mathbb{P}(Z_1+\cdots+Z_{n-1}=\ell,\, Z_n=k-\ell) \\
&= \sum_{\ell=0}^{k}\mathbb{P}(Z_1+\cdots+Z_{n-1}=\ell)\mathbb{P}(Z_n=k-\ell) \\
&\qquad \text{（$Z_1+\cdots+Z_{n-1}$ と Z_n の独立性より）} \\
&= \sum_{\ell=0}^{k}\frac{(n-1)^{\ell}}{\ell!}e^{-(n-1)}\frac{1}{(k-\ell)!}e^{-1} \\
&= \frac{e^{-n}}{k!}\sum_{\ell=0}^{k}\frac{k!}{\ell!(k-\ell)!}(n-1)^{\ell}1^{k-\ell} \\
&= \frac{e^{-n}}{k!}\{(n-1)+1\}^k \\
&\quad \text{（二項定理 $(a+b)^k=\sum_{\ell=0}^{k}a^{\ell}b^{k-\ell}$ から）} \\
&= \frac{n^k}{k!}e^{-n}
\end{aligned}$$

となるので，$Z_1+\cdots+Z_n$ は母数 n のポアソン分布に従うことがわかる．このことを繰り返し利用すれば，$Z_1+\cdots+Z_n$ は母数 n のポアソン分布に従うので，X_n と同じ分布になる．

すなわち，X_n は母数 1 のポアソン分布に独立同一に従う確率変数の和 $Z_1 + Z_2 + \cdots + Z_n$ と同じ分布をもつ．$\bar{Z}_n = \frac{1}{n}(Z_1 + \cdots + Z_n)$ とおく．したがって，任意の $x \in \mathbb{R}$ に対して，

$$
\begin{aligned}
\mathbb{P}\left(\frac{X_n - n}{\sqrt{n}} \leq x\right) &= \mathbb{P}\left(\frac{Z_1 + \cdots + Z_n - n}{\sqrt{n}} \leq x\right) \\
&= \mathbb{P}\left(\frac{\frac{1}{n}(Z_1 + \cdots + Z_n) - 1}{\frac{1}{n}\sqrt{n}} \leq x\right) \\
&= \mathbb{P}\left(\frac{\sqrt{n}\{\bar{Z}_n - \mathbb{E}[Z_1]\}}{\sqrt{\mathbb{VAR}[Z_1]}} \leq x\right) \\
&\xrightarrow{d} N(0,\, 1)
\end{aligned}
$$

となることがわかる．$Z_1,\, Z_2,\, \ldots,\, Z_n$ は独立同一に母数 1 のポアソン分布に従うことに注意して，最後のところで中心極限定理を用いた．