

順序統計量を用いた指數分布の特徴づけ

統計数理研究所 柳本武美

指數分布は待ち行列、寿命試験等でよく用いられる分布である。これは指數分布の記憶忘却の分布と呼ばれるようすと特徴的な性質をもつことによる。

指數分布の特徴づけは正規分布の特徴づけとみて程度平行的に結果が得られる。もう少し具体的に言うと、正規分布の特徴づけでは二つの独立な確率変数 X_1, X_2 の convolution $Y_1 = \sin\theta X_1 + \cos\theta X_2$ と $Y_2 = \cos\theta X_1 - \sin\theta X_2$ を考えて、その独立性あるいは Y_1 と Y_2 が同一分布に従う性質から X_1, X_2 が正規であることを示す。指數分布の特徴づけでは順序統計量を用いる。

X_1, \dots, X_m を分布関数が $F(x)$ である確率変数 X からの大きさのサンプルとする。 $X_{(1)}, \dots, X_{(m)}$ を（小さい方から）順序統計量とする。 $F(x) = 1 - e^{-\lambda x}$ ($\lambda > 0$) のとき、即ち X が位置母数 λ の指數分布に従うとき、 $nX_{(1)} - (m-1)(X_{(2)} - X_{(1)})$

, ---, $(X_m - X_{m-1})$ は互いに独立で同一分布 \times 従う, その分布関数は元の $F(x)$ と一致する。逆にこの性質は指数分布と特徴づけられる。

本稿では $m = 2$ の場合について新しい結果を含めて紹介する。この問題の周辺には興味ある問題が山積してあるが, 正規分布・特徴づけにおける特性関数のような有力な手段も見出せないまま, 解かれていよい。

以下では $X \sim Y$ は X と Y が同一分布 \times 従うことを, $X \perp Y$ は X と Y が独立であることを表す。

$$(1) X_{(2)} - X_{(1)} \sim X$$

Puri ら[4]の結果を少し拡張して次の結果を得る。

定理 1. $X_{(2)} - X_{(1)} \sim X$ であれば, 次の 4 つの性質のうちいずれか一つを満足する。

$$\text{i)} F(x) = 1 - e^{-\lambda x} (\lambda > 0)$$

$$\text{ii)} \text{ある } T > 0 \text{ に対して } P(X=kT) = 2p_0(1-p_0)(1-2p_0)^{k-1} \\ \text{for } k \geq 1$$

$$P(X=0) = p_0$$

$$\text{iii)} \text{ある } \tau > 0 \text{ に対して, } P(X=\tau) = P(X=0) = 1/2$$

$$\text{iv)} P(X=0) = 0 \quad (\text{これは自明の場合で, 以降にこの場合を除く})$$

証明. [4] で殆んど証明された通り完結されればよい。残る、 x の場合に極端連続でなくて有界でなく、 $x' > x > 0$ で $F(x') = F(x)$ が成り立たない場合である。この場合も X は任意次のモノメントとなる。

state space を $[0, \infty)$ と transition probability を $K(x, \Gamma)$ とし、

$$K(x, \Gamma_x) = \int_{y \in \Gamma_x^+} (dF(x+y) + dF(x-y)) + \int_{y \in \Gamma_x^-} dF(x+y)$$

$$+ \int_{y \in \Gamma_x^0} dF(x+y)$$

$\therefore C \Gamma_x^+ = \{y \mid x+y > 0, y \in \Gamma\}, \Gamma_x^- = \{y \mid x-y, y \in \Gamma\}, \Gamma^0 = \{0\}$ である。条件が $dF(x)$ は stationary probability である。すなはち密度関数が

$$f(x) = 1 - F(x-0) \quad \text{for } x > 0$$

$$= 1/2 \quad \text{for } x = 0$$

で表わされる確率分布も stationary である。従って上で定義したマルコフチェインが ergodic であることを示せばよい。この為 K. I. Feller [3] の p266 定理2を用いれば容易く示せる。

(2) $2X_{n_1} \sim X$

この条件は分布関数を用いて表わすと

$$(1 - F(\frac{x}{2}))^2 = 1 - F(x) \text{ for any } x$$

と一致する。 $X > 0$ に退化した場合を除くと, $P(X > 0) = 1$ 。

だから, $\Psi(y)$ を

$$\Psi(y) = \log(-\log(1 - F(e^y))) - y$$

と定義すると上の条件は $\Psi(y)$ の周期 $\log 2$ をもつことと一致する。従って, $2X_{(1)} \sim X$ は指數分布と特徴づけない。

$X_{(1),m}$ を大きさ m の標本の最小順序統計量とするとき、次の定理が成り立つ。

定理. $\log m / \log 2$ が無理数とする。 $2X_{(1)} \sim mX_{(1),m} \sim X$ であれば、 $F(x) = 1 - e^{-\lambda x}$ ($\lambda > 0$) である。

系. [2] すべての n に対して、 $nX_{(1),n} \sim X$ が成り立つは $F(x) = 1 - e^{-\lambda x}$ ($\lambda > 0$) である。

(3) $2X_{(1)} \sim X_{(2)} - X_{(1)}$

open problem である。

(4) $2X_{(1)} \leq X_{(2)} - X_{(1)}$

この条件については多くの結果が得られた。様々な拡張もなされた。[1] で上の条件を満たすのは指數分布や幾何分布のみずれいかであることを示している。

Schanbhag [5] は平均の存在を仮定して、 $E(X_{(2)} - X_{(1)})$

$| X_{(1)} = x$) が一定である時の条件を求めて。 $F(x)$ が連續であれば指数分布であり、マスホイニトが適当なくて、 $\mu = \lambda + 1$ で $\{ \lambda t + \mu \mid t \text{ は整数} \}$ に含まれる場合は幾何分布である。

この条件は [6] の記号を使うと $(2X_{(1)}, X_{(2)} - X_{(1)}) \in \mathcal{P}(3, E)$ へ $\mathcal{M}(3, E)$ と殆んど同じになる。一般にはこの条件を満たす分布は数多く存在する。

上では平均を扱ったが、 $X_{(2)} - X_{(1)} | X_{(1)} = x$ と X との間に、

$P(X_{(2)} - X_{(1)} | X_{(1)} = x > X) = P(X_{(2)} - X_{(1)} | X_{(1)} < X)$

が成り立つ場合を考える。ただし X_1, X_2, X は互に独立であるとする。この条件はやはり [6] の記号を使うと $(2X_{(1)}, X_{(2)} - X_{(1)}) \in \mathcal{P}(3, 0) \cap \mathcal{M}(3, 0)$ で表わされる。次の定理は最初の定理の系として得られる。

定理. $F(x)$ が連續とする。 $(2X_{(1)}, X_{(2)} - X_{(1)}) \in \mathcal{P}(3, 0) \cap \mathcal{M}(3, 0)$ であれば、 $F(x) = 1 - e^{-\lambda x}$ ($\lambda > 0$) である。

REFERENCES

- [1] Crawford, G. B. (1966). Characterization of geometric and exponential distributions, Ann. Math. Statist., 37, 1790-1795.
- [2] Desu, M. M. (1971). A characterization of the exponential distribution by order statistics, Ann. Math. Statist., 42, 837-838.
- [3] Feller, W. (1966). An Introduction to Probability Theory and Its Applications, Wiley, New York.

- [4] Puri, P. S. and Rubin, H. (1970). A characterization based on the absolute difference of two i. i. d. random variables, Ann. Math. Statist., 41, 2113-2122.
- [5] Shanbhag, D. N. (1970). The characterizations for exponential and geometric distributions, J. Amer. Statist. Assoc., 65, 1256-1259.
- [6] Yanagimoto, T. (1972). Families of positively dependent random variables, Ann. Inst. Math. Statist., 24, 559-573.