4.3 順序統計量

定義 4.5 $X_1, X_2, \dots, X_n$ をランダム標本としたとき，これを小さい順に並べかえたものを

$$X_{(1)}, X_{(2)}, \dots, X_{(n)}$$

を記し，これらを**順序統計量**という．すなわち，

$$\begin{aligned} X_{(1)} &= \min\{X_1, X_2, \dots, X_n\}, \\ X_{(2)} &= X_1, X_2, \dots, X_n \text{ の中で 2 番目に小さいもの}, \\ &\vdots \\ X_{(n)} &= \max\{X_1, X_2, \dots, X_n\}, \end{aligned}$$

である．

定理 4.7 $X_1, X_2, \dots, X_n$ を離散型分布 $f_X(x_i) = p_i, i = 1, 2, \dots,$ からの標本の大きさが n のランダム標本とする．ただし，$x_1 < x_2 < \cdots$ は X の台(4-3) とする．さらに，$P_i = \sum_{j=1}^{i} p_j, i = 1, 2, \dots, P_0 = 0$ とおく．$X_{(1)}, X_{(2)}, \dots, X_{(n)}$ を標本の順序統計量としたとき，

$$P(X_{(j)} \leq x_i) = \sum_{k=j}^{n} \binom{n}{k} P_i^k (1 - P_i)^{n-k}, \qquad j = 1, 2, \dots, n, i = 1, 2, \dots, \tag{4.9}$$

と

$$P(X_{(j)} = x_i) = \sum_{k=j}^{n} \binom{n}{k} [P_i^k (1 - P_i)^{n-k} - P_{i-1}^k (1 - P_{i-1})^{n-k}], \tag{4.10}$$

となる．

証明 $i \in \{1, 2, \dots\}$ を固定し，Y を確率変数とし，

$$Y = \#\{X_j, j = 1, 2, \dots, n | X_j \leq x_i\}$$

とする．ここで，確率変数 Z_j を $\{X_j \leq x_i\}$ が起こったとき，$Z_j = 1$，さもなければ，$Z_j = 0$ と定める．$X_1, X_2, \dots, X_n$ は同一分布に従うので，各 j について，

$$\mathbb{P}(Z_j = 1) = P_i = \mathbb{P}(X_j \leq x_i)$$

である．また，$Z_1, Z_2, \dots, Z_n$ は独立である．さらに，$Y = \sum_{j=1}^{n} Z_j$ に注意すれば，Y は母数 n, P_i の二項分布 $Bin(n, P_i)$ に従うことがわかる．

事象 $\{X_{(j)} \leq x_i\}$ は事象 $\{Y \geq j\}$ と同じなので，

$$\mathbb{P}(X_{(j)} \leq x_i) = \mathbb{P}(Y \geq j)$$

となり，(4.9) は示された．(4.10) を示すためには，

$$\mathbb{P}(X_{(j)} = x_i) = \mathbb{P}(X_{(j)} \leq x_i) - \mathbb{P}(X_{(j)} \leq x_{i-1})$$

を考えればよい．$i = 1$ の場合は $\mathbb{P}(X_{(j)} = x_i) = X_{(j)} \leq x_i)$ よりわかる．以上から (4.10) を示された． □

定理 4.8 $X_1, X_2, \ldots, X_n$ を分布関数 $F_X(x)$ と確率密度関数 $f_X(x)$ をもつ連続型分布からの標本の大きさが n のランダム標本とする．$X_{(1)}, X_{(2)}, \ldots, X_{(n)}$ を標本の順序統計量としたとき，$X_{(j)}, j = 1, 2, \ldots, n,$ の確率密度関数は

$$f_{X_{(j)}}(x) = \frac{n!}{(j-1)!(n-j)!} f_X(x)[F_X(x)]^{j-1}[1-F_X(x)]^{n-j}, \tag{4.11}$$

となる．

証明 $X_{(j)}$ の分布関数を求め，その導関数を計算することにより，$X_{(j)}$ の確率密度関数を求めることにする．

実数 x を固定する．Y を確率変数とし，

$$Y = \#\{X_j, j = 1, 2, \ldots, n | X_j \leq x\}$$

とする．ここで，確率変数 Z_j を $\{X_j \leq x\}$ が起こったとき，$Z_j = 1$，さもなければ，$Z_j = 0$ と定める．$X_1, X_2, \ldots, X_n$ は同一分布に従うので，各 j について，

$$\mathbb{P}(Z_j = 1) = P_i = F_X(x)$$

である．また，$Z_1, Z_2, \ldots, Z_n$ は独立である．さらに，$Y = \sum_{j=1}^n Z_j$ に注意すれば，Y は母数 n, P_i の二項分布 $Bin(n, F_X(x))$ に従うことがわかる．これらより

$$F_{X_{(j)}}(x) = \mathbb{P}(Y \geq j) = \sum_{k=j}^{n} \binom{n}{k} [F_X(x)]^k [1-F_X(x)]^{n-k}$$

となる．したがって，$X_{(j)}$ の確率密度関数は

$$\begin{aligned}
f_{X_{(j)}}(x) &= \frac{d}{dx} F_{X_{(j)}}(x) \\
&= \sum_{k=j}^{n} \binom{n}{k} \Big(k[F_X(x)]^{k-1}[1-F_X(x)]^{n-k} f_X(x) \\
&\quad -(n-k)[F_X(x)]^k[1-F_X(x)]^{n-k-1} f_X(x) \Big) \\
&= \binom{n}{j} j f_X(x)[F_X(x)]^{j-1}[1-F_X(x)]^{n-j} \\
&\quad + \sum_{k=j+1}^{n} \binom{n}{k} k[F_X(x)]^{k-1}[1-F_X(x)]^{n-k} f_X(x) \\
&\quad - \sum_{k=j}^{n-1} \binom{n}{k} (n-k)[F_X(x)]^k[1-F_X(x)]^{n-k-1} f_X(x) \\
&= \frac{n!}{(j-1)!(n-j)!} f_X(x)[F_X(x)]^{j-1}[1-F_X(x)]^{n-j} \\
&\quad + \sum_{k=j}^{n-1} \binom{n}{k+1} (k+1)[F_X(x)]^k[1-F_X(x)]^{n-k-1} f_X(x) \\
&\quad - \sum_{k=j}^{n-1} \binom{n}{k} (n-k)[F_X(x)]^k[1-F_X(x)]^{n-k-1} f_X(x)
\end{aligned}$$

となる．最後に，

$$\binom{n}{k+1}(k+1) = \frac{n!}{k!(n-k-1)!} = \binom{n}{k}(n-k)$$

から (4.11) は示される． □

例 4.4 $X_1, X_2, \ldots, X_n$ を $(0, 1)$ 上の一様分布からの標本の大きさが n のランダム標本とする．すなわち，

$$
\begin{aligned}
F_X(x) &= \begin{cases} x, & 0 < x < 1, \\ 0, & x \leq 0, \\ 1, & x \geq 1 \end{cases}, \\
f_X(x) &= \begin{cases} 1, & 0 < x < 1, \\ 0, & \text{その他}, \end{cases}
\end{aligned}
$$

である．このとき，(4.11) から

$$
f_{X_{(j)}} = \begin{cases} \frac{n!}{(j-1)!(n-j)!} x^{j-1}(1-x)^{n-j}, & 0 < x < 1, \\ 0, & \text{その他} \end{cases}, \qquad j = 1, 2, \ldots, n
$$

となる．また，補遺の代表的な広義積分 (iii) から

$$
\begin{aligned}
\mathbb{E}[X_{(j)}] &= \frac{n!}{(j-1)!(n-j)!} \int_0^1 x^j (1-x)^{n-j}, dx \\
&= \frac{n!}{(j-1)!(n-j)!} \frac{\Gamma(j+1)\Gamma(n-j+1)}{\Gamma(n+2)} \\
&= \frac{n!}{(j-1)!(n-j)!} \frac{j!(n-j)!}{(n+1)!} \\
&= \frac{j}{n+1}
\end{aligned}
$$

となる．

定理 4.9 $X_1, X_2, \ldots, X_n$ を分布関数 $F_X(x)$ と確率密度関数 $f_X(x)$ をもつ連続型分布からの標本の大きさが n のランダム標本とする．$X_{(1)}, X_{(2)}, \ldots, X_{(n)}$ を標本の順序統計量としたとき，$X_{(i)}$ と $X_{(j)}, 1 \leq i \leq j \leq n$，の同時確率密度関数は

$$
\begin{aligned}
&f_{X_{(i)}, X_{(j)}}(u, v) \\
&= \begin{cases} \frac{n!}{(i-1)!(j-1-i)!(n-j)!} f_X(u) f_X(v) [F_X(u)]^{i-1} \\ \qquad \times [F_X(v) - F_X(v)]^{j-i-1} [1 - F_X(x)]^{n-j}, & -\infty < u < v < \infty, \\ 0, & \text{その他} \end{cases}
\end{aligned}
\tag{4.12}
$$

となる．

証明 証明は省略する． □

系 4.2 $X_{(1)}$ と $X_{(n)}$ の同時確率密度関数は

$$
f_{X_{(1)}, X_{(n)}}(x_1, x_n) = \begin{cases} n(n-1)[F_X(x_n) - F_X(x_1)]^{n-2} f_X(x_1) f(x_n) & x_1 < x_n, \\ 0 & x_1 \geq x_n \end{cases}
$$

で与えられる．

例 4.5 $X_1, X_2, \ldots, X_n, n \geq 2$ は独立同一に $(0, 1)$ 上の一様分布に従うとする．すなわち，各 $X_i, i = 1, 2, \ldots, n$ は確率密度関数と分布関数

$$
f_X(x) = \begin{cases} 1 & 0 < x < 1, \\ 0 & \text{（その他）} \end{cases}, \qquad F_X(x) = \begin{cases} 0 & x \leq 0, \\ x & 0 < x < 1, \\ 1 & x \geq 1 \end{cases}
$$

を持つ．定理 4.8 から

$$f_{X_{(j)}}(x) = \begin{cases} \frac{n!}{(j-1)!(n-j)!} x^{j-1}(1-x)^{n-1} & 0 < x < 1, \\ 0 & \text{その他）} \end{cases}$$

となる．