第 4 章「多変量確率ベクトルの分布」

1．同時分布関数（Joint Distribution Function）

・**同時分布関数** n 個の確率変数 $X_1, \cdots, X_n$ に対して，

$$F(x_1, \cdot, x_n) = P(X_1 \le x_1, \cdots, X_n \le x_n)$$

を $(X_1, \cdots, X_n)$ の同時分布関数という．

・**周辺分布関数**

$$F_i(x_i) = F(\infty, \cdots, \infty, x_i, \infty, \cdots, \infty) = P(X_i \le x_i), \qquad i = 1, \cdots, n$$

を，X_i の周辺分布関数という．

・**同時密度関数** 連続的な場合には，同時密度関数

$$f(x_1, \cdot, x_n) = \frac{\partial^n}{\partial x_1 \cdots \partial x_n} F(x_1, \cdot, x_n)$$

が定義される．

2．確率ベクトルの期待値と分散

以下，X, Y は n 次元確率ベクトル，a, b は n 次元定数ベクトルとする．これらは，列ベクトルで定義される．すなわち，

$$X = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}, \qquad Y = \begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix}, \qquad a = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}, \qquad b = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

また，A を $n \times p$ の行列，B を $m \times q$ の行列とする．すなわち，

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{np} \end{pmatrix}, \qquad B = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1q} \\ b_{21} & b_{22} & \dots & b_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \dots & b_{mq} \end{pmatrix}$$

・転置行列

行列 A の転置行列を A' で表す．列ベクトル X の転置 X' は行ベクトルとなる．$X'X$ はスカラーで X の要素の 2 乗和となる．

・平均ベクトル

$$E(X) = \begin{pmatrix} E(X_1) \\ \vdots \\ E(X_n) \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix} = \mu_X$$

・分散共分散行列

$$V(X) = E\left[(X-\mu_X)(X-\mu_X)'\right] = \begin{pmatrix} Cov(X_1,X_1) & Cov(X_1,X_2) & \ldots & Cov(X_1,X_n) \\ Cov(X_2,X_1) & Cov(X_2,X_2) & \cdots & Cov(X_2,X_n) \\ \vdots & \vdots & \ddots & \vdots \\ Cov(X_n,X_1) & Cov(X_n,X_2) & \ldots & Cov(X_n,X_n) \end{pmatrix} = \Sigma$$

（問１） $E(XX') = \Sigma + \mu_X\mu_X'$ となることを示せ.

・線形結合のモーメント

$$E(a'X) = a'E(X) = a'\mu_X, \qquad V(a'X) = a'V(X)a = a'\Sigma a$$

（問２） $X_1,\cdots,X_n$ が互いに独立で，同一分布に従うとき，標本平均 $\bar{X}$ に対して，次のことが成り立つことを示せ.

$$E(\bar{X}) = E(X_1), \qquad V(\bar{X}) = \frac{1}{n}V(X_1)$$

・線形変換のモーメント

$$E(A'X) = A'E(X) = A'\mu_X, \qquad V(A'X) = A'V(X)A = A'\Sigma A$$

・2 次形式の期待値 C を $n\times n$ の対称行列とするとき，

$$E(X'CX) = \mathrm{tr}\,(C\Sigma) + \mu_X'C\mu_X$$

（注） スカラーの場合は，$E(cX^2) = c(\sigma^2+\mu_X^2)$.

・2 つの確率ベクトルの共分散

$$Cov(X,Y) = E\left[(X-\mu_X)(Y-\mu_Y)'\right] = \begin{pmatrix} Cov(X_1,Y_1) & Cov(X_1,Y_2) & \ldots & Cov(X_1,Y_m) \\ Cov(X_2,Y_1) & Cov(X_2,Y_2) & \cdots & Cov(X_2,Y_m) \\ \vdots & \vdots & \ddots & \vdots \\ Cov(X_n,Y_1) & Cov(X_n,Y_2) & \ldots & Cov(X_n,Y_m) \end{pmatrix}$$

３．多次元確率分布の例

・多項分布 $M_N(n;\ p,q)$

k 通りの事象 $A_1,\cdots,A_k$ のいずれかが確率 $P(A_i)=p_i$ で起きる独立試行を n 回行うものとする．このとき，それぞれの事象が起きる回数 $X_1,\cdots,X_k$ の同時確率分布は

$$P(X_1=x_1,\cdots,X_k=x_k) = \frac{n!}{x_1!\,\cdots\,x_k!}\,p_1^{x_1}\,\cdots\,p_k^{x_k}$$

で与えられる．この分布を**多項分布**といい，$M_N(n;\,p_1,\cdots,p_k)$ で表す.

・共分散と相関係数

$$Cov(X_i, X_j) = -np_i p_j, \qquad Corr(X_i, X_j) = -\left(\frac{p_i}{1-p_i}\,\frac{p_j}{1-p_j}\right)^{1/2}$$

・順序統計量への応用

$X_1, \cdots, X_n$ を独立，同一分布に従う確率変数とするとき，これらを小さい方から並べてできるもの $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$ を**順序統計量**という． $X_1, \cdots, X_n$ が独立，同一分布に従っても，$X_{(1)}, X_{(2)}, \cdots, X_{(n)}$ は独立でなく，同一分布にも従わない．

・順序統計量の周辺分布：確率変数 X_r の分布関数を $F(x)$ とするとき，$X_{(r)}$ の分布関数 $F_r(x)$ と，密度関数（存在する場合）は次のように求めることができる．

$$\begin{aligned} F_r(x) &= P\left(X_{(r)} \leq x\right) = \sum_{i=r}^{n} P\left(X_{(i)} \leq x < X_{(i+1)}\right) \\ &= \sum_{i=r}^{n} \binom{n}{i} (F(x))^i \, (1-F(x))^{n-i} \\ f_r(x) &= \frac{d\,F_r(x)}{d\,x} = r\binom{n}{r} (F(x))^{r-1} \, (1-F(x))^{n-r} \, f(x) \end{aligned}$$

（例） [0, 1] 上の一様分布の場合

$$f_r(x) = r\binom{n}{r} x^{r-1} (1-x)^{n-r}$$

・順序統計量の同時密度：順序統計量 $X_{(r)}$ と $X_{(s)}$ $(r < s)$ の同時密度関数（存在を仮定）$f_{rs}(x, y)$ は多項分布の確率を考えることにより，次のように求めることができる．

$$f_{rs}(x, y) = \frac{n!}{(r-1)!\,(s-r-1)!\,(n-s)!} (F(x))^{r-1} \, f(x) \, (F(y)-F(x))^{s-r-1} \, f(y) \, (1-F(y))^{n-s}$$

・多変量正規分布

n 次元確率ベクトル X の密度関数が

$$f(x) = \frac{1}{\sqrt{2\pi}^n \, |\Sigma|^{1/2}} \exp\left\{-\frac{1}{2}(x-\mu)' \, \Sigma^{-1} \, (x-\mu)\right\}$$

と表されるとき，X は n 次元正規分布に従うといい，$X \sim N_n(\mu, \Sigma)$ と表す．ここで，

$$\mu = E(X), \qquad \Sigma = V(X) = E\left[(X-\mu)(X-\mu)'\right] : \text{正値対称}$$

n **次元多変量正規分布の性質**

1. $X \sim N_n(\mu, \Sigma)$ ならば，$Z = \Sigma^{-1/2}\,(X-\mu) \sim N_n(0, I_n)$

2. $Z \sim N_n(0, I_n)$ ならば， $X = \mu + \Sigma^{1/2}\,Z \sim N_n(\mu, \Sigma)$

3. $X_2|X_1 \sim N\left(\mu_2 + \Sigma_{21}\Sigma_{11}^{-1}(X_1-\mu_1), \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma 12\right)$
ここで，

$$X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}, \quad \mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \quad \begin{pmatrix} \Sigma_{11} & \Sigma 12 \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$$

4. X_1 と X_2 が多変量正規で互いに独立 $\Leftrightarrow Cov(X_1, X_2) = \Sigma_{12} = 0$

（注） 正値対称行列 Σ に対して，平方根行列 $\Sigma^{1/2}$ は次のように定義される．

$P\Sigma P' = \Lambda$ （P は直交行列，Λ は対角要素がすべて正の対角行列）

$\Rightarrow \Sigma = P'\Lambda P = P'\Lambda^{1/2}\Lambda^{1/2}P = P'\Lambda^{1/2}PP'\Lambda^{1/2}P$

$\Rightarrow \Sigma^{1/2} = P'\Lambda^{1/2}P$

• n **次元多変量正規分布の積率母関数**

まず，$Z \sim N(0, I_n)$ のとき，

$$\begin{aligned} M_Z(t) &= E\left(\exp(t'Z)\right) = E\left(\exp(t_1 Z_1 + \cdots + t_n Z_n)\right) \\ &= \prod_{j=1}^{n} E\left(\exp(t_j Z_j)\right) = \prod_{j=1}^{n} \exp\left(\frac{1}{2}t_j^2\right) \\ &= \exp\left(\frac{1}{2}t't\right) \end{aligned}$$

が成り立つから，$X = \mu + \Sigma^{1/2}Z \sim N(\mu, \Sigma)$ に対しては

$$\begin{aligned} M_X(t) &= E\left(\exp(t'X)\right) = E\left(\exp(t'(\mu + \Sigma^{1/2}Z))\right) \\ &= \exp(t'\mu)E\left(\exp(t'\Sigma^{1/2}Z)\right) = \exp(t'\mu)E\left(\exp\left(\left(\Sigma^{1/2}t\right)'Z\right)\right) \\ &= \exp(t'\mu)\exp\left(\frac{1}{2}\left(\Sigma^{1/2}t\right)'\Sigma^{1/2}t\right) \\ &= \exp\left(t'\mu + \frac{1}{2}t'\Sigma t\right) \end{aligned}$$

（問） 積率母関数を使って，$X \sim N(\mu, \Sigma)$ のとき，$E(X_1X_2X_3)$ を求めよ．

４．大数の法則（Law of Large Numbers: LLN）

定理（大数の弱法則: WLLN）確率変数の列 $X_1, X_2, \cdots, X_n$ が互いに独立に平均 μ，分散 σ^2 の同一分布に従うとき，標本平均 $\bar{X}_n$ は μ に**確率収束**する．すなわち，任意の正数 ε に対して

$$\lim_{n\to\infty} P\left(|\bar{X}_n - \mu| \geq \varepsilon\right) = 0$$

[証明]　チェビシェフの不等式を使う．

５．中心極限定理（Central Limit Theorem: CLT）

定理: CLT 確率変数の列 $X_1, X_2, \cdots, X_n$ が互いに独立に平均 μ，分散 σ^2 の同一分布に従うとき，標本平均 $\bar{X}_n$ を標準化した量は N(0,1) に分布収束する．すなわち，

$$Z_n = \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} = \frac{\sum_{j=1}^n (X_j - \mu)}{\sqrt{n}\sigma} \quad \longrightarrow \quad N(0,1)$$

[証明] Z_n の特性関数が N(0,1) の特性関数に収束することをいえばよい．まず，X_j を標準化した量 $Y_j = (X_j - \mu)/\sigma$ の特性関数

$$\phi(t) = E\left(\exp\left(it\,\frac{X_j - \mu}{\sigma}\right)\right) = E(\exp(itY_j))$$

を定義する．このとき，Z_n の特性関数は

$$\begin{aligned}
\phi_n(t) &= E(\exp(itZ_n)) = E\left(\exp\left(it\frac{\sum_{j=1}^n (X_j - \mu)}{\sqrt{n}\sigma}\right)\right) \\
&= E\left(\exp\left(it\frac{\sum_{j=1}^n Y_j}{\sqrt{n}}\right)\right) = \prod_{j=1}^n E\left(\exp\left(\frac{itY_j}{\sqrt{n}}\right)\right) \\
&= \left\{\phi\left(\frac{t}{\sqrt{n}}\right)\right\}^n
\end{aligned}$$

となる．ここで，$E(Y_j) = 0,\ E(Y_j^2) = 1$ に注意して，テーラー展開を使うことにより

$$\begin{aligned}
\phi\left(\frac{t}{\sqrt{n}}\right) &= E\left(\exp\left(\frac{itY_j}{\sqrt{n}}\right)\right) \\
&= E\left[1 + \frac{itY_j}{\sqrt{n}} + \frac{1}{2!}\left(\frac{itY_j}{\sqrt{n}}\right)^2 + \cdots\right] \\
&= 1 - \frac{t^2}{2n} + o\left(\frac{1}{n}\right)
\end{aligned}$$

を得る（この点に関する厳密な証明は，例えば Shiryayev, A.N. の *Probability*, Springer-Verlag を見よ）．このことから，

$$\begin{aligned}
\phi_n(t) &= \phi\left(\frac{t}{\sqrt{n}}\right) = \left(1 - \frac{t^2}{2n} + o\left(\frac{1}{n}\right)\right)^n \\
&\to \exp\left(-\frac{t^2}{2}\right)
\end{aligned}$$

となり，N(0,1) の特性関数が得られる．

6．正規近似（Normal Approximation）

・正規近似： $X_1, X_2, \cdots, X_n$ が互いに独立に平均 μ, 分散 σ^2 の同一分布に従うとき，和 $X_1 + X_2 + \cdots + X_n$ がほぼ正規分布に従うこと（CLT）を使って，和に関連した確率の近似計算を行うこと．

例１：一様分布の場合　X_1, X_2, X_3, X_4 が互いに独立に [10,20] 上の一様分布に従うとき，和 $Y = X_1 + X_2 + X_3 + X_4$ が 70 以上となる確率の近似値を求める．

[解] まず，$E(X_1) = 15$, $V(X_1) = (20-10)^2/12 = 100/12$ であるから，

$$E(Y) = E(X_1 + X_2 + X_3 + X_4) = 4 \times E(X_1) = 60, \qquad V(Y) = 4 \times V(X_1) = \frac{100}{3}$$

となる．したがって，$Z = (Y-60)/\sqrt{100/3}$ がほぼ N(0,1) であることを使って，

$$\begin{aligned} P(Y \geq 70) &= P\left(Z \geq \frac{70-60}{\sqrt{100/3}}\right) = P(Z \geq \sqrt{3}) \\ &\cong 0.042 \end{aligned}$$

を得る（正確な値は $1/24 = 0.042$ である）．

例２：二項分布の場合 成功の確率が 0.3 の独立試行において，20 回のうち 7 回以上成功する確率の近似値を求める．

[解] $X \sim B_N(n,p)$ ならば，$Z = (X - np)/\sqrt{np(1-p)}$ がほぼ N(0,1) に従うから，$X \sim B_N(20, 0.3)$ のとき，

$$\begin{aligned} P(X \geq 7) &= P\left(Z \geq \frac{7-6}{\sqrt{20 \times 0.3 \times 0.7}}\right) = P(Z \geq 0.488) \\ &\cong 0.31 \end{aligned}$$

を得る．正確な値は

$$P(X \geq 7) = \sum_{k=7}^{20} {}_{20}C_k \, 0.3^k \, 0.7^{20-k} = 0.392$$

となるので，若干の乖離がある．

・不連続補正 離散分布の場合の近似を改善するための補正であり，二項分布の場合には次のように補正する．$X \sim B_N(n,p)$ のとき，

$$P(X \geq j) = P(X \geq j - 0.5), \qquad P(X \leq k) = P(X \leq k + 0.5)$$

・例２の場合

$$\begin{aligned} P(X \geq 7) &= P(X \geq 6.5) = P\left(Z \geq \frac{6.5-6}{\sqrt{4.2}}\right) = P(Z \geq 0.244) \\ &= 0.404 \end{aligned}$$