6　同時確率分布

X, Y を離散型確率変数とする。これを (X, Y) と表し、2 次元確率変数と呼ぶ。

$$(X,Y) \text{ の値域} = \Big\{(x_i, y_j) \mid i = 1, 2, \cdots, M;\ j = 1, 2, \cdots, N\Big\} \tag{1}$$

とする。(X, Y) の性質は、

$$P\Big((X,Y) = (x_i, y_j)\Big) = P\Big(X = x_i, Y = y_j\Big) = f(x_i, y_j) \quad (i, j = 1, 2, \cdots) \tag{2}$$

によって定まる。これを (X, Y) の同時確率分布と言う。勿論のこと、$\sum_{i=1}^{M} \sum_{j=1}^{N} f(x_i, y_j) = 1$ が成り立つ。

記号として、

$$E(X) = \mu_X,\ E(Y) = \mu_Y,\ V(X) = \sigma_X^2,\ V(Y) = \sigma_Y^2,\ C(X,Y) = \sigma_{XY} \tag{3}$$

を用いる。

6.1　簡単な例

2 次元確率変数 (X, Y) の同時確率分布が次のように与えられているとする。

$Y \backslash X$	0	1	2	$P(Y = y)$
0	1/8	3/8	0	4/8
1	0	1/8	3/8	4/8
$P(X = x)$	1/8	4/8	3/8	1

(i) X の周辺分布 $P(X = x)$ は次の通り。

x	0	1	2
$P(X = x)$	$\frac{1}{8}$	$\frac{4}{8}$	$\frac{3}{8}$

(ii) Y の周辺分布 $P(Y = y)$ は次の通り。

y	0	1
$P(Y = y)$	$\frac{1}{2}$	$\frac{1}{2}$

(iii) $E(X)$ と $V(X)$ とを求める。周辺分布から計算すればよい：

$$E(X) = \sum_{x=0}^{2} x\ P(X = x) = 0 \times \frac{1}{8} + 1 \times \frac{4}{8} + 2 \times \frac{3}{8} = \frac{5}{4} \tag{4}$$

$$E(X^2) = \sum_{x=0}^{2} x^2\ P(X = x) = 0^2 \times \frac{1}{8} + 1^2 \times \frac{4}{8} + 2^2 \times \frac{3}{8} = 2 \tag{5}$$

$$V(X) = E(X^2) - \{E(X)\}^2 = 2 - (5/4)^2 = \frac{7}{16} \tag{6}$$

(iv) $E(Y)$ と $V(Y)$ とを求める。上と同様に出来る。$E(Y) = 1/2$、$V(Y) = 1/4$。
(v) $E(XY)$ を求める。これは同時分布から求める。

$$\begin{aligned} E(XY) &= \sum_{x=0}^{2}\sum_{y=0}^{1} xy\ P(X = x,\ Y = y) \\ &= (0 \times 0 \times (1/8) + (0 \times 1 \times (3/8) + (0 \times 2 \times 0) + (1 \times 0 \times 0) + (1 \times 1 \times (1/8)) + (1 \times 2 \times (3/8) = 7/8. \end{aligned}$$

(vi) 共分散 $\mathrm{C}(X, Y)$ を求めよう。公式: $\mathrm{C}(X, Y) = E(XY) - E(X)E(Y)$ を用いる。

$$\mathrm{C}(X, Y) = E(XY) - E(X)E(Y) = 7/8 - (5/4) \times (1/2) = 1/4. \tag{7}$$

(vii) 相関係数 $\rho(X, Y) = \mathrm{C}(X, Y)/\sqrt{V(X)V(Y)}$ を求める：

$$\rho(X, Y) = \frac{\mathrm{C}(X, Y)}{\sqrt{V(X)V(Y)}} = \frac{\frac{1}{4}}{\sqrt{\frac{7}{16} \times \frac{1}{4}}} \approx 0.756 \tag{8}$$

以下、条件付き分布について考察する。

1. $Y = 0$ が与えられた時の X の条件付分布を求める：

$$\begin{aligned} P(X = 0|Y = 0) &= \frac{P(X = 0, Y = 0)}{P(Y = 0)} = \frac{1/8}{4/8} = \frac{1}{4}, \\ P(X = 1|Y = 0) &= \frac{P(X = 1, Y = 0)}{P(Y = 0)} = \frac{3/8}{4/8} = \frac{3}{4}, \\ P(X = 2|Y = 0) &= \frac{P(X = 2, Y = 0)}{P(Y = 0)} = \frac{0}{4/8} = 0. \end{aligned} \tag{9}$$

2. $Y = 0$ が与えられた時の X の条件付平均と条件付分散を求める：上記の分布を使って平均と分散を計算すればよい。

$$E(X|Y = 0) = \sum_{x=0}^{1} x\ P(X = x|Y = 0) = 0 \times (1/4) + 1 \times (3/4) + 2 \times 0 = 3/4, \tag{10}$$

$$\begin{aligned} V(X|Y = 0) &= E(X^2|Y = 1) - \{E(X|Y = 1)\}^2 \\ &= \{0^2 \times (1/4) + 1^2 \times (3/4) + 2^2 \times 0\} - (3/4)^2 = 3/16. \end{aligned} \tag{11}$$

3. $Y = 1$ が与えられた時の X の条件付平均を求める：$Y = 1$ が与えられた時の条件付分布は、

$$\begin{aligned} P(X = 0|Y = 1) &= \frac{P(X = 0, Y = 1)}{P(Y = 1)} = \frac{0}{4/8} = 0, \\ P(X = 1|Y = 1) &= \frac{P(X = 1, Y = 1)}{P(Y = 1)} = \frac{1/8}{4/8} = 1/4. \\ P(X = 2|Y = 1) &= \frac{P(X = 2, Y = 1)}{P(Y = 1)} = \frac{3/8}{4/8} = 3/4 \end{aligned} \tag{12}$$

であるから、$E(X|Y = 1) = 0 \times 0 + 1 \times (1/4) + 2 \times (3/4) = 7/4.$.

4. 期待値の繰り返しの公式を用いて $E(X)$ を求める。

$$E(X) = E(X|Y = 0)\ P(Y = 0) + E(X|Y = 1)\ P(Y = 1) = (3/4) \times (1/2) + (7/4) \times (1/2) = 5/4.$$

6.2 独立性と無相関性

1. (無相関だが独立ではない例) 2 次元確率変数 (X, Y) の同時確率分布が次の通りに与えられているとする。

$X \backslash Y$	-1	0	1
-1	0.1	0.1	0.1
0	0.2	0	0.2
1	0.1	0.1	0.1

   (a) X の周辺分布を求め、$E(X)$ と $V(X)$ を計算せよ。

   (b) Y の周辺分布を求め、$E(Y)$ と $V(Y)$ を計算せよ。

   (c) 共分散 $C(X, Y)$ と相関係数 ρ_{XY} とを計算せよ。

   (d) X と Y が独立ではないことを示せ。

2. (計算練習) 四角形の縦の長さと横の長さを計り、それぞれ Xcm と Ycm という結果を得るものとする。X と Y は独立とする。$X \sim U(0, 1)$、$Y \sim U(0, 1)$ であるとする。このとき四角形の面積 XY の平均 $E(XY)$ と分散 $V(XY)$ を求めよ。

6.3 基本定理

定理 1. $X_1, \cdots, X_n$ は互いに独立に同一の確率分布 F に従うものとする。(以後、この仮定を**独立同一分布の仮定**と呼び、$X_1, X_2, \cdots, X_n \overset{\perp\!\!\!\perp}{\sim} F$ と書くことにする。)

$$E(X_1) = \cdots = E(X_n) \equiv \mu, \quad V(X_1) = \cdots = V(X_n) \equiv \sigma^2$$

と置く。このとき、$X_1, \cdots, X_n$ の平均 $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ は次の 2 式を満足する。

$$E(\bar{X}) = \mu, \quad V(\bar{X}) = \frac{\sigma^2}{n}. \tag{13}$$

例 6.1. (典型例) n 個の確率変数 $X_1, \cdots, X_n$ が互いに独立に同一の分布 F に従っているとする．分布 F の平均は μ、分散は σ^2 であるとする。$T = \sum_{i=1}^{n} X_i$ とおく．このとき，次の結果が得られる．

1. ベルヌーイ分布 $Ber(p)$ の場合。この場合、$\mu = p$，$\sigma^2 = p(1-p)$ であるから，

$$E(\bar{X}) = p, V(\bar{X}) = \frac{p(1-p)}{n} \tag{14}$$
$$E(T) = np, V(T) = np(1-p) \tag{15}$$

2. ポアソン分布 $P_O(\lambda)$ の場合。この場合、$\mu = \lambda$，$\sigma^2 = \lambda$ であるから，

$$E(\bar{X}) = \lambda, V(\bar{X}) = \frac{\lambda}{n} \tag{16}$$
$$E(T) = n\lambda, V(T) = n\lambda \tag{17}$$

3. 正規分布 $N(\mu, \sigma^2)$ の場合。この場合、

$$E(\bar{X}) = \mu, V(\bar{X}) = \frac{\sigma^2}{n} \tag{18}$$
$$E(T) = n\mu, V(T) = n\sigma^2 \tag{19}$$

4. 指数分布 $E_X(\lambda)$ の場合。この場合、$\mu = 1/\lambda$、$\sigma^2 = 1/\lambda^2$ であるから、

$$E(\bar{X}) = 1/\lambda, V(\bar{X}) = \frac{1}{n\lambda^2} \tag{20}$$

その他の確率分布についても同様である.

例 6.2. （応用例）上式とチェビシェフの不等式を組み合わせて、$\bar{X}$ のばらつきを調べよう。

チェビシェフの不等式を復習すると次の通り：確率変数 Z は、平均 $E(Z) = \alpha$、分散 $V(Z) = \beta^2$ であるとする。このとき、

$$P(\alpha - k\beta \leq Z \leq \alpha + k\beta) \geq 1 - 1/k^2.$$

1. あるテレビ番組の視聴率が 20%であったとする（通常この数値は未知だが、今は全知の立場に立つことにする）。400 世帯に視聴の有無について尋ねるとする。回答を X_i $(i = 1, 2, \cdots, 400)$ と表し、視聴していたならば $X_i = 1$、そうでなければ $X_i = 0$ と記録するとする。このとき、$X_1, \cdots, X_{400}$ は互いに独立に同一のベルヌーイ分布 $Ber(0.2)$ に従っていると見なしてよい。このとき、$\bar{X} = (X_1 + X_2 + \cdots + X_{400})/400$ はどれくらいばらつくか。

   上の例の 1. において $p = 0.2$ として、

   $$E(\bar{X}) = 0.2, \quad V(\bar{X}) = \frac{0.2 \times 0.8}{400} = 0.0004, \ D(\bar{X}) = \sqrt{0.0004} = 0.02. \tag{21}$$

   従って、チェビシェフの不等式より、

   $$P(0.2 - 2 \times 0.02 \leq \bar{X} \leq 0.2 + 2 \times 0.02) = P(0.16 \leq \bar{X} \leq 0.24) \geq 1 - 1/2^2 = 3/4 \tag{22}$$

   となる。次回示す通り、n が大きいときは元の分布が何であっても、$\bar{X}$ の分布は正規分布で近似することができる（中心極限定理）から、$\bar{X} \sim N(0.2, (0.02)^2)$ が近似的に成り立つ。従って、上式の確率は 95.4%となる。

2. ある測定機器の測定誤差 ϵ は正規分布 $N(0, 2^2)$ であるとする。ある測定対象の真の重量を μ とする。従って、測定結果 $X = \mu + \epsilon$ の分布は $N(\mu, 2^2)$ となる。この測定機器を用いて、同じ測定対象を独立に 100 回測定し、その結果 $X_1, \ldots, X_{100}$ の平均 $\bar{X} = \frac{1}{100}\sum_{i=1}^{100} X_i$ を計算するものとする。上の例の 3. において $\sigma^2 = 2^2$ として、

   $$E(\bar{X}) = \mu, \quad V(\bar{X}) = \frac{2^2}{100} = 0.04, \ D(\bar{X}) = \sqrt{0.04} = 0.2. \tag{23}$$

   従って、チェビシェフの不等式より、

   $$P(\mu - 2 \times 0.2 \leq \bar{X} \leq \mu + 2 \times 0.2) \geq 1 - 1/2^2 = 3/4 \tag{24}$$

   となる。次回示す通り、元の分布が正規分布の場合、$\bar{X}$ も正規分布に従うので $\bar{X} \sim N(100, (0.2)^2)$ が成り立つ。従って、上式の確率は 95.4%となる。今の場合は、上式のように表現するよりも、

   $$P(\mu - 2 \times 0.2 \leq \bar{X} \leq \mu + 2 \times 0.2) = P(\bar{X} - 2 \times 0.2 \leq \mu \leq \bar{X} + 2 \times 0.2) \geq 1 - 1/2^2 = 3/4 \tag{25}$$

   と変形した方が応用性がある。例えば、$\bar{X} = 24$ であったとすれば、区間 [23.6, 24.4] なる区間が得られ、これは未知の μ を確率 3/4 以上で（正確には 95.4%で）含む区間の実現値である。