第2章　期待値の基礎事項

この章では, 確率変数の期待値と関連するものを導入し, その基本的な性質を述べる. 節 2.1 では, 確率変数に対する期待値を定義する. 節 2.2 では, 期待値を確率ベクトルに対して定義する. 節 2.3 では, 分布のばらつきを測る量である分散と共分散を導入し, 基本的な性質を説明する. 節 2.4 では, 条件付き分布に対する期待値を導入する. 節 2.5 では, 分布を特定する量である積率母関数を導入し, その基本的な性質を説明する.

2.1　期待値

X を確率空間 $(\Omega, \mathcal{A}, \mathsf{Pr})$ 上の確率変数とする. X が離散型のとき, その p.m.f. を p とし, X の取り得る値を $x_1, x_2, \ldots$ とする. 連続型のとき, その p.d.f. も p と書くことにする.

定義 2.1. $g : \mathbb{R} \to \mathbb{R}$ を可測[1] 関数とする. 確率変数 $g(X)$ の期待値 $\mathsf{E}[g(X)]$ を次のように定義する.

(1) $g(x) \geq 0\,(x \in \mathbb{R})$ のとき

$$\mathsf{E}[g(X)] = \begin{cases} \displaystyle\sum_{n=1}^{\infty} g(x_n)\mathsf{p}(x_n) & (\text{離散型の場合}) \\ \displaystyle\int_{-\infty}^{\infty} g(x)\mathsf{p}(x)\,\mathrm{d}x & (\text{連続型の場合}) \end{cases}$$

と定義する. 右辺は ∞ を許せば, 必ず存在する.

(2) 一般の可測関数 g に対して

$$g^+(x) = \max\{g(x), 0\}, \quad g^-(x) = \max\{-g(x), 0\}$$

と定義すれば, $g^+(x) \geq 0$, $g^-(x) \geq 0\,(x \in \mathbb{R})$ となる. $\mathsf{E}[g^+(X)]$ または $\mathsf{E}[g^-(X)]$ のいずれかが有限ならば

$$\mathsf{E}[g(X)] := \mathsf{E}[g^+(X)] - \mathsf{E}[g^-(X)]$$

[1] 可測関数の定義は定義 **??** を参照.

と定義する．$\mathsf{E}[g^+(X)] = \mathsf{E}[g^-(X)] = \infty$ のときは, $g(X)$ の期待値は定義されない．$\mathsf{E}[g^+(X)] < \infty$ かつ $\mathsf{E}[g^-(X)] < \infty$ のとき, $\mathsf{E}[g(X)]$ は有限となる．

補題 2.2. X を確率変数とする．関数 $g, h : \mathbb{R} \to \mathbb{R}$ を可測とする．

(1) $h(X)$ は有限の期待値を持つ[2]とする．$0 \le g(x) \le h(x)\ (\forall x \in \mathbb{R})$ ならば

$$\mathsf{E}[g(X)] \le \mathsf{E}[h(X)]$$

となる．

(2) $\mathsf{E}[|g(X)|] < \infty$, $\mathsf{E}[|h(X)|] < \infty$ を仮定する．$a, b \in \mathbb{R}$ に対して

$$\mathsf{E}[ag(X) + bh(X)] = a\,\mathsf{E}[g(X)] + b\,\mathsf{E}[h(X)]$$

となる．

Proof. (1) の証明: $h(x) - g(x) \ge 0\ (x \in \mathbb{R})$ であるので, 定義 2.1(1) と積分の性質から

$$\mathsf{E}\big[h(X) - g(X)\big] = \int_{-\infty}^{\infty} \underbrace{\big\{h(x) - g(x)\big\}\mathsf{p}^X(x)}_{\ge 0}\,\mathrm{d}x \ge 0$$

となる．さらに, (1) を用いると $\mathsf{E}[h(X)] \ge \mathsf{E}[g(X)]$ がわかる．

(2) の証明: 確率変数 X が p.d.f. p^X を持つ場合を示す．$\big|ag(x)+bh(x)\big| \le |a||g(x)| + |b||h(x)|$ と $g(X)$ と $h(X)$ の期待値が有限なので, (1) から $ag(X) + bh(X)$ は有限な期待値をもつ．このことから, 定義 2.1(1) のステップは省略できる．あとは, 積分の線型性から

$$\begin{aligned}
\mathsf{E}\big[ag(X) + bh(X)\big] &= \int_{-\infty}^{\infty} \big\{ag(x) + bh(x)\big\}\mathsf{p}^X(x)\,\mathrm{d}x \\
&= a\int_{-\infty}^{\infty} g(x)\mathsf{p}^X(x)\,\mathrm{d}x + b\int_{-\infty}^{\infty} h(x)\mathsf{p}^X(x)\,\mathrm{d}x \\
&= a\mathsf{E}[g(X)] + b\mathsf{E}[h(X)]
\end{aligned}$$

がわかる． □

補題 2.3. X を非負値確率変数とする．このとき,

$$\mathsf{E}[X] = 0 \Rightarrow \mathsf{Pr}(X = 0) = 1$$

となる．

[2] g, h を共に非負値関数としているので, $\mathsf{E}[|h(X)|] < \infty$ だけを仮定すればよい．もし, g, h を可測関数とすると $\mathsf{E}[|g(X)|] < \infty$ の仮定も必要になる．なぜならば, $g(x) \le h(x)\ (x \in \mathbb{R})$ と $\mathsf{E}[|h(X)|] < \infty$ だけだと $\mathsf{E}[h(X)]$ が定義できない場合があるからである．なので, 非負値関数に関する主張とした方のが実用的である．

Proof. まず

$$
\begin{aligned}
A &:= \bigl\{\omega \in \Omega;\ X(\omega) > 0\bigr\} = \{X > 0\},\\
A_n &:= \left\{\omega \in \Omega;\ X(\omega) > \frac{1}{n}\right\} \qquad (n = 1,\, 2, \ldots)
\end{aligned}
$$

とおく．すると $A_1 \subset A_2 \subset \cdots$ かつ $\displaystyle\bigcup_{n=1}^{\infty} A_n = A$ となる．よって，補題 1.7(5) から

$$
\lim_{n\to\infty} \mathsf{Pr}(A_n) = \mathsf{Pr}\left(\bigcup_{n=1}^{\infty} A_n\right) = \mathsf{Pr}(A) \tag{2.1}
$$

となる．一方，$A_n \subset A$ なので，$\mathbb{1}_{A_n}(\omega) \le \mathbb{1}_A(\omega)\,(x \ge 0)\,(\omega \in \Omega)$ で[3]ある．このことに注意すると

$$
\begin{aligned}
0 = \mathsf{E}[X] &= \mathsf{E}\bigl[\underbrace{X\mathbb{1}_{A^c}}_{\ge 0} + X\mathbb{1}_A\bigr] = \mathsf{E}\bigl[X\mathbb{1}_A\bigr] \ge \mathsf{E}\bigl[X\mathbb{1}_{A_n}\bigr]\\
&\ge \frac{1}{n}\mathsf{E}[\mathbb{1}_{A_n}] = \frac{1}{n}\mathsf{Pr}(A_n)
\end{aligned}
$$

を得る[4]．したがって，$\mathsf{Pr}(A_n) = 0\,(\forall n \in \mathbb{N})$ がわかる．(2.1) と合わせると $\mathsf{Pr}(A) = 0$ となる．よって

$$
\begin{aligned}
1 = \mathsf{Pr}(A^c) &= \mathsf{Pr}\Bigl(\bigl(\{X \le 0\} \cap \{X \ge 0\}\bigr) \cup \bigl(\{X \le 0\} \cap \{X < 0\}\bigr)\Bigr)\\
&= \mathsf{Pr}(X = 0) + \underbrace{\mathsf{Pr}(X < 0)}_{=0} = \mathsf{Pr}(X = 0)
\end{aligned}
$$

が示せた． □

注意 2.4. $0 < q < r$ に対して

$$
\mathsf{E}[|X|^r] < \infty \Rightarrow \mathsf{E}[|X|^q] < \infty
$$

となる．

Proof. Young の不等式 (系 3.13)

$$
ab \le \frac{a^s}{s} + \frac{b^t}{t};\quad a,\, b > 0;\quad s,\, t > 0;\quad \frac{1}{t} + \frac{1}{s} = 1
$$

---

[3] $\mathbb{1}_A$ は $\mathbb{1}_A : \Omega \ni \omega \mapsto \mathbb{1}_A(\omega) = \begin{cases} 1 & (\omega \in A) \\ 0 & (\omega \in A^c) \end{cases}$ なる確率変数である．

[4]

において

$$a = |x|^q; s = \frac{r}{q}\,(0 < s < 1) \quad b = 1; t = \frac{1}{1 - s/2}\,(0 < t < 1)$$

とおくと

$$|x|^q \leq \frac{|x|^r}{s} + \frac{1}{t}$$

を得る. よって, 補題 2.2(1) からわかる. □

問 2.1. 連続確率変数 X が p.d.f. p をもつとき, 補題 2.4 の証明を具体的に書け. $q < r$ なので, $|x|^{q-r} \leq 1\,(|x| > 1)$ となることを使えばよい.

定義 2.5. (1) $k = 1, 2, \ldots$ に対して, $\mathsf{E}[|X|^k] < \infty$ のとき, $\mathsf{E}[X^k]$ を X の k **次モーメント** (または**積率**) という.

(2) $\mathsf{E}[|X|] < \infty$ のとき $\mathsf{E}[X]$ を X **の平均値**[5] という.

(3) $\mathsf{E}[X^2] < \infty$ のとき X **の分散**を

$$\mathsf{Var}[X] := \mathsf{E}[\{X - \mathsf{E}[X]\}^2]$$

で定義する.

(4) 部分集合 $A \subset \mathbb{R}$ に対して

$$\mathbb{1}_A(x) = \begin{cases} 1 & (x \in A) \\ 0 & (x \notin A) \end{cases}$$

を A の**指示関数**[6] という. $A \in \mathcal{B}(\mathbb{R})$ のとき

$$\mathsf{E}[\mathbb{1}_A(X)] = \mathsf{Pr}(X \in A)$$

となる.

注意 2.6. 分散 $\mathsf{Var}[X]$ は X の分布の平均 μ まわりの散らばりを測る量である. 分散がおおきいほど分布は広がっていることになる. □

注意 2.7. (1) X は確率空間 $(\Omega, \mathcal{A}, \mathsf{Pr})$ 上の離散型確率変数とする. p^X を X の p.m.f. とし

$$S_X := \{x \in \mathbb{R};\, \mathsf{p}^X(x) > 0\}$$

[5]簡単に「平均」ともいう.
[6]指示関数は $\mathbb{R}$ の任意の部分集合に定義ができることに注意をせよ.

とする. 関数 $g : \mathbb{R} \to \mathbb{R}$ に対して

$$Y = g(X)$$

とおく. $y \in \mathbb{R}$ に対して

$$A_y := \{x \in S_X;\, g(x) = y\}$$

とおき

$$S_Y := \{y \in \mathbb{R};\, A_y \neq \varnothing\}$$

とする. $y \in S_Y$ に対して

$$\mathsf{p}^Y(y) := \mathsf{Pr}(Y = y) = \mathsf{Pr}(X \in A_y) = \sum_{x \in A_y} \mathsf{Pr}(X = x) = \sum_{x \in A_y} \mathsf{p}^X(x)$$

と書ける. 正項級数は項の順番を並び替えてもその値は変わらないので

$$\begin{aligned}\sum_{y \in S_Y} |y| \mathsf{p}^Y(y) &= \sum_{y \in S_Y} |y| \sum_{x \in A_y} \mathsf{p}^X(x) = \sum_{y \in S_Y} \sum_{x \in A_y} |g(x)| \mathsf{p}^X(x) \\ &= \sum_{x \in S_X} |g(x)| \mathsf{p}^X(x)\end{aligned}$$

となる. さらにいずれかの和が有限ならば

$$\sum_{y \in S_Y} y \mathsf{p}^Y(y) = \sum_{y \in S_Y} y \sum_{x \in A_y} \mathsf{p}^X(x) = \sum_{y \in S_Y} \sum_{x \in A_y} g(x) \mathsf{p}^X(x) = \sum_{x \in S_X} g(x) \mathsf{p}^X(x)$$

となる. よって

$$\mathsf{E}[Y] = \mathsf{E}[g(X)]$$

となる.
(2) 確率変数 X は連続型とする. p^X を X の p.d.f. とし

$$S_X := \{x \in \mathbb{R};\, \mathsf{p}^X(x) > 0\}$$

とする. 関数 $g : S_X \to \mathbb{R}$ は狭義単調増加かつ C^1 級とする. このとき

$$Y = g(X), \quad S_Y := \{y \in \mathbb{R};\, y = g(x)\, (\exists x \in S_X)\}$$

とする. ここで $g : S_X \to S_Y$ と制限[7] すれば g の逆関数 $g^{-1} : S_Y \to S_X$ が存在し

$$\begin{aligned}\mathsf{F}^Y(y) &:= \mathsf{Pr}(Y \leq y) = \mathsf{Pr}\big(g^{-1}(Y) \leq g^{-1}(y)\big) = \mathsf{Pr}(X \leq g^{-1}(y)) \\ &= \mathsf{F}^X(g^{-1}(y))\end{aligned}$$

[7]制限したものを同じ g を用いて表現するのは, 記号の乱用である. 記号が煩雑になるので, 記号を乱用した.

となる．さらに $g(g^{-1}(y)) = y$ より

$$\frac{\mathrm{d}}{\mathrm{d}y} g^{-1}(y) = \frac{1}{\overset{\bullet}{g}\,(g^{-1}(y))}$$

となる．ただし $\overset{\bullet}{g}\,(y) = \dfrac{\mathrm{d}g}{\mathrm{d}y}(y)$ である．これらより，$y \in S_Y$ に対して

$$\begin{aligned}\mathsf{p}^Y(y) &= \frac{\mathrm{d}}{\mathrm{d}y}\mathsf{F}^Y(y) = \frac{\mathrm{d}}{\mathrm{d}y}\mathsf{F}^X(g^{-1}(y)) = \overset{\bullet}{\mathsf{F}}{}^X\,(g^{-1}(y))\frac{\mathrm{d}}{\mathrm{d}y}g^{-1}(y) \\ &= \mathsf{p}^X(g^{-1}(y))\frac{1}{\overset{\bullet}{g}\,\big(g^{-1}(y)\big)}\end{aligned}$$

となる．よって

$$\mathsf{p}^Y(y) = \mathsf{p}^X\big(g^{-1}(y)\big)\frac{1}{\overset{\bullet}{g}\,(g^{-1}(y))}$$

となる．g が狭義単調減少の場合もふくめると

$$\mathsf{p}^Y(y) = \mathsf{p}^X\big(g^{-1}(y)\big)\frac{1}{\big|\,\overset{\bullet}{g}\,\big(g^{-1}(y)\big)\big|}$$

となる．

以上のことから

$$\begin{aligned}\mathsf{E}[Y] &= \int_{S_Y} y\mathsf{p}^Y(y)\,\mathrm{d}y \\ &= \int_{S_Y} y\mathsf{p}^X(g^{-1}(y))\frac{1}{\overset{\bullet}{g}\,\big(g^{-1}(y)\big)}\,\mathrm{d}y \\ &= \int_{S_X} g(x)\mathsf{p}^X(x)\,\mathrm{d}x \qquad \big(\because x = g^{-1}(y) \Leftrightarrow y = g(x);\ \mathrm{d}y = \overset{\bullet}{g}\,(x)\mathrm{d}x\big) \\ &= \mathsf{E}[g(X)]\end{aligned}$$

がわかる．　□

2.2　確率ベクトルの期待値

確率変数 X, Y を確率空間 $(\Omega, \mathcal{A}, \mathsf{Pr})$ 上で定義された確率変数とする．(X, Y) を確率ベクトルという．

確率ベクトル (X, Y) が離散型のときその同時 p.m.f. を $\mathsf{p}(x, y)$ とし，連続型のときその同時 p.d.f. も $\mathsf{p}(x, y)$ と書くことにする．

定義 2.8. 可測関数 $g : \mathbb{R}^2 \to \mathbb{R}$ に対して $g(X, Y)$ の期待値を次のように定義する．

(1) $g \geq 0$ のとき

$$\mathsf{E}[g(X,\,Y)] = \begin{cases} \displaystyle\sum_{x,\,y} g(x,\,y)\mathsf{p}(x,\,y) & (\text{離散型の場合}) \\ \displaystyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,\,y)\mathsf{p}(x,\,y)\,\mathrm{d}x\mathrm{d}y & (\text{連続型の場合}) \end{cases}$$

で定義する.
(2) 一般の g に対して

$$g^{+}(x,\,y) := \max\{g(x,\,y),\,0\}, \quad g^{-}(x,\,y) := \max\{-g(x,\,y),\,0\}$$

と定義すれば $g^{+},\,g^{-} \geq 0$ となる. $\mathsf{E}[g^{+}(X,\,Y)]$ または $\mathsf{E}[g^{-}(X,\,Y)]$ のいずれかが有限ならば

$$\mathsf{E}[g(X,\,Y)] := \mathsf{E}[g^{+}(X,\,Y)] - \mathsf{E}[g^{-}(X,\,Y)]$$

で定義する. $\mathsf{E}[g^{+}(X,\,Y)] = \mathsf{E}[g^{-}(X,\,Y)] = \infty$ のときは, $g(X,\,Y)$ の期待値は定義されない. $\mathsf{E}[g^{+}(X,\,Y)] < \infty$ かつ $\mathsf{E}[g^{-}(X,\,Y)] < \infty$ のとき, $\mathsf{E}[g(X,\,Y)]$ は有限の値を取る.

注意 2.9. 3 つ以上の確率変数 $X_1,\,X_2,\,\ldots,\,X_n$ に対しても期待値を定義 2.8 と同様に定義する. □

定理 2.10. $X_1,\,X_2,\,\ldots,\,X_n$ を確率変数とし, 各 $X_j\,(j = 1,\,2,\,\ldots,\,n)$ の期待値は有限とする. $a_1,\,a_2,\,\ldots,\,a_n$ を定数としたとき

$$\mathsf{E}\Bigl[\sum_{j=1}^{n} a_j X_j\Bigr] = \sum_{j=1}^{n} a_j \mathsf{E}[X_j]$$

となる.

Proof. まず, $\bigl|\sum_{j=1}^{n} a_j x_j\bigr| \leq \sum_{j=1}^{n} a_j |x_j|$ なので, 補題 2.2(1) から $\sum_{j=1}^{n} a_j X_j$ は有限な期待値をもつ. $\boldsymbol{X} := (X_1,\,X_2,\,\ldots,\,X_n)^{\top}$ とおき, $\boldsymbol{X}$ は同時 p.d.f. $\mathsf{p}^{\boldsymbol{X}}$ を持つとする. すると定義 2.8 から 補題 2.2(2) と同様の期待値の線型性から

$$\begin{aligned} \mathsf{E}\Bigl[\sum_{j=1}^{n} a_j X_j\Bigr] &= \int_{\mathbb{R}^n} \Bigl\{\sum_{j=1}^{n} a_j x_j\Bigr\} \mathsf{p}^{\boldsymbol{X}}(x_1,\,x_2,\,\ldots,\,x_n)\,\mathrm{d}x_1\mathrm{d}x_2\cdots\mathrm{d}x_n \\ &= \sum_{j=1}^{n} a_j \int_{\mathbb{R}^n} x_j \mathsf{p}^{\boldsymbol{X}}(x_1,\,x_2,\,\ldots,\,x_n)\,\mathrm{d}x_1\mathrm{d}x_2\cdots\mathrm{d}x_n \\ &= \sum_{j=1}^{n} a_j \mathsf{E}\bigl[X_j\bigr] \end{aligned}$$

がわかる. □

例 2.11. $0 < p < 1$ と $j = 1, 2, \ldots, n$ に対して, $X_j \sim \mathsf{Ber}(p)$ とする. このとき

$$\mathsf{E}[X_j] = \sum_{x=0}^{1} x\mathsf{Pr}(X = x) = 0 \times (1-p) + 1 \times p = p$$

である. $S = \sum_{j=1}^{n} X_j$ としたとき

$$\mathsf{E}[S] = \sum_{j=1}^{n} \mathsf{E}[X_j] = np$$

がわかる. □

2.3 分散と共分散

X を確率変数とし, $\mathsf{E}[X^2] < \infty$ とする. X の分散を

$$\mathsf{Var}[X] := \mathsf{E}[(X-\mu)^2]$$

で定義した. ただし $\mu = \mathsf{E}[X]$ と書いた. さらに $\sqrt{\mathsf{Var}[X]}$ を X の標準偏差という.

定理 2.12. 以下の確率変数は有限の 2 次の積率を持つとする. このとき, 次が成立する.

(1) $\mathsf{Var}[X] = \mathsf{E}[X^2] - \{\mathsf{E}[X]\}^2$ となる.

(2) 定数 $a, b \in \mathbb{R}$ に対して

$$\mathsf{Var}[aX + b] = a^2\mathsf{Var}[X]$$

となる.

(3) X と Y は独立で $\mathsf{E}[|XY|] < \infty$ とする. このとき

$$\mathsf{E}[XY] = \mathsf{E}[X]\mathsf{E}[Y]$$

となる.

(4) $X_1, X_2, \ldots, X_n$ は独立とし, $\mathsf{E}[X_j^2] < \infty \, (j = 1, 2, \ldots, n)$ とする. $a_1, a_2, \ldots, a_n$ は定数としたとき

$$\mathsf{Var}[a_1X_1 + a_2X_2 + \cdots + a_nX_n] = a_1^2\mathsf{Var}[X_1] + a_2^2\mathsf{Var}[X_2] + \cdots + a_n^2\mathsf{Var}[X_n]$$

となる.

Proof. (1), (2) は分散を期待値の記号を用いて表現し, 期待値の線型性を用いて計算すればよい. (3) については, 連続型の場合を示す. (X, Y) の同時 p.d.f. は, X と Y の周辺 p.d.f. p^X と p^Y を用いて $\mathsf{p}^X(x)\mathsf{p}^Y(y)$ という形でかけるので

$$\begin{aligned}\mathsf{E}[XY] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\mathsf{p}^X(x)\mathsf{p}^Y(y)\,\mathrm{d}x\mathrm{d}y = \int_{-\infty}^{\infty} x\mathsf{p}^X(x)\,\mathrm{d}x \int_{-\infty}^{\infty} y\mathsf{p}^Y(y)\,\mathrm{d}y \\ &= \mathsf{E}[X]\mathsf{E}[Y]\end{aligned}$$

となること [8] がわかる. (4) は, 分散を期待値の記号を用いて表現し, (3) に注意する. 期待値の線型性を用いて計算すればよい. □

例 2.13. (例 2.11 の続き) 例 2.11 の設定に加えて $X_1, X_2, \ldots, X_n$ は互いに独立とする. 定理 2.10 に注意すれば

$$\begin{aligned}\mathsf{Var}[X_j] &= \mathsf{E}[X_j^2] - \{\mathsf{E}[X_j]\}^2 = \sum_{x=0}^{1} x^2 \mathsf{Pr}(X = x) - p^2 \\ &= 0 \times (1-p) + 1 \times p - p^2 = p(1-p)\end{aligned}$$

となる. これと定理 2.12(4) から

$$\mathsf{Var}[S] = \sum_{j=1}^{n} \mathsf{Var}[X_j] = np(1-p)$$

がわかる. □

定理 2.14. $n \geq 2$ とし, $X_1, X_2, \ldots, X_n$ は i.i.d. 確率変数列とし

$$\mathsf{E}[X_1] = \mu, \quad \mathsf{Var}[X_1] = \sigma^2$$

とする. ただし, $\mu \in \mathbb{R}$, $0 < \sigma < \infty$ である. $X_1, X_2, \ldots, X_n$ に基づく**標本平均**を

$$\overline{X}_n = \frac{1}{n}\sum_{j=1}^{n} X_j$$

で定義し, **標本 (不偏) 分散**を

$$S_n^2 = \frac{1}{n-1}\sum_{j=1}^{n}\bigl(X_j - \overline{X}_n\bigr)^2$$

で定義する. このとき

$$(1)\ \mathsf{E}\bigl[\overline{X}_n\bigr] = \mu; \qquad (2)\ \mathsf{Var}\bigl[\overline{X}_n\bigr] = \frac{\sigma^2}{n}; \qquad (3)\ \mathsf{E}\bigl[S_n^2\bigr] = \sigma^2$$

となる.

[8] $\mathsf{Var}[X] < \infty$, $\mathsf{Var}[Y] < \infty$ なので, Cauchy-Schwarz の不等式から $\mathsf{E}[|XY|] < \infty$ が確認できるので, Fubini 定理から積分順序の入れ替えが保証されることがわかる.

Proof. (1), (2) は定理 2.10, 2.12(4) よりわかる. (3) を証明するために

$$\sum_{j=1}^{n}\big(X_j-\overline{X}_n\big)^2=\sum_{j=1}^{n}(X_j-\mu)^2-n\big(\overline{X}_n-\mu\big)^2$$

に注意する. 定理 2.12(1) より

$$\mathsf{E}\Bigg[\sum_{j=1}^{n}(X_j-\mu)^2\Bigg]=n\sigma^2$$

がわかる. さらに, (1) と (2) より

$$\mathsf{E}\big[\big(\overline{X}_n-\mu\big)^2\big]=\mathsf{Var}\big[\overline{X}_n\big]=\frac{\sigma^2}{n}$$

が示される. したがって

$$\begin{aligned}\mathsf{E}[S_n^2]&=\frac{1}{n-1}\Bigg\{\sum_{j=1}^{n}\mathsf{E}\big[(X_j-\mu)^2\big]-n\mathsf{E}\big[\big(\overline{X}_n-\mu\big)^2\big]\Bigg\}\\&=\frac{1}{n-1}\big\{n\sigma^2-\sigma^2\big\}=\sigma^2\end{aligned}$$

がわかる. □

定義 2.15. X と Y は確率変数とし

$$\mathsf{E}[X]=\mu_X,\quad \mathsf{Var}[X]=\sigma_X^2,\quad \mathsf{E}[Y]=\mu_Y,\quad \mathsf{Var}[Y]=\sigma_Y^2$$

とする. ただし $\mu_X,\,\mu_Y\in\mathbb{R}$, $0<\sigma_X,\,\sigma_Y<\infty$ とする. このとき X と Y の共分散を

$$\mathsf{Cov}[X,\,Y]=\mathsf{E}\big[(X-\mu_X)(Y-\mu_Y)\big]$$

で定義し, X と Y の (Pearson) の相関係数を

$$\rho:=\rho[X,\,Y]:=\frac{\mathsf{Cov}[X,\,Y]}{\sigma_X\,\sigma_Y}$$

で定義する.

定理 2.16. X, Y, Z は 2 次の積率が有限な確率変数とする.
(1) 共分散は

$$\mathsf{Cov}[X,\,Y]=\mathsf{E}[XY]-\mathsf{E}[X]\mathsf{E}[Y]$$

と書き直せる.
(2) $\mathsf{Cov}[X,\,Y]=\mathsf{Cov}[Y,\,X]$ である.
(3) 定数 a, b に対して

$$\mathsf{Cov}[aX+bY,\,Z]=a\mathsf{Cov}[X,\,Z]+b\mathsf{Cov}[Y,\,Z]$$

となる.
(4) 相関係数は

$$-1 \leq \rho[X, Y] \leq 1$$

をみたす.
(5) ある定数 $a, b\,(a \neq 0)$ が存在して $Y = aX + b$ となったとき

$$a > 0 \Rightarrow \rho[X, Y] = 1,$$
$$a < 0 \Rightarrow \rho[X, Y] = -1$$

である.
(6) X と Y が独立のとき

$$\mathsf{Cov}[X, Y] = 0$$

となる.

Proof. (1) は期待値の線型性からわかる. (2)(3) は共分散の定義と期待値の線型性よりわかる. (4) は $\mathsf{Var}[X] = \mathsf{Var}[Y] = 0$ のときは明らかである. よって $\mathsf{Var}[X] \neq 0$ として証明する. 補題 2.2(2) より

$$\mathsf{Var}[X]t^2 - 2\mathsf{Cov}[X, Y]t + \mathsf{Var}[Y] = \mathsf{E}[\{t(X - \mathsf{E}[X]) + Y - \mathsf{E}[Y]\}^2] \geq 0$$

となる. ここで, 判別式をとれば

$$\{\mathsf{Cov}[X, Y]\}^2 + \mathsf{Var}[X]\mathsf{Var}[Y] \leq 0 \tag{2.2}$$

がわかる. (5) は (2), (3) および定理 2.12(2) からわかる. (6) は定理 2.12(3) からわかる. □

注意 2.17. (2.2) において, $\mathsf{E}[X] = \mathsf{E}[Y] = 0$ の場合を考えると Cauchy-Schwarz の不等式

$$\left|\mathsf{E}[XY]\right| \leq \sqrt{\mathsf{E}[X^2]}\sqrt{\mathsf{E}[Y^2]}$$

と呼ばれる重要なものを得る.

注意 2.18. (6) の逆は一般に正しくない. 反例は下の問いをみよ. □

問 2.2. Z を $[0, 2\pi]$ 上の一様分布とし

$$X = \cos Z, \quad Y = \sin Z$$

とおく.

(1) $\mathsf{E}[X] = 0$, $\mathsf{E}[Y] = 0$, $\mathsf{E}[XY] = 0$ を確かめよ.

(2) $\mathsf{Pr}(X \le 1/2,\, Y \le 1/2) = \mathsf{Pr}((\pi/3) \le Z \le (5/6)\pi)$ を確かめよ.

(3) $\mathsf{Pr}(X \le 1/2) = \mathsf{Pr}((\pi/3) \le Z \le (7/3)\pi)$ を確かめよ[9].

定理 2.19. (1) $X,\, Y$ は有限な 2 次の積率を持つ確率変数とする. このとき $\mathsf{Var}[X+Y] = \mathsf{Var}[X] + \mathsf{Var}[Y] + 2\mathsf{Cov}[X,\, Y]$ となる.
(2) $d \ge 2$ とする. $X_1,\, X_2,,\, \ldots,\, X_d$ は有限な 2 次の積率を持つ確率変数とする. このとき

$$\mathsf{Var}\left[\sum_{j=1}^{d} a_j X_j\right] = \sum_{j=1}^{d} a_j^2 \mathsf{Var}[X_j] + 2\sum_{j=1}^{d}\sum_{\ell=j+1}^{d} a_j a_\ell \mathsf{Cov}[X_j,\, X_\ell]$$

となる.

Proof. 共分散を期待値で表現し, 展開して期待値の線型性を用いればよい. □

問 2.3. 定理 2.19(1) の証明を具体的に書け.

定義 2.20. $X_1,\, X_2,\, \ldots,\, X_d$ を有限な 2 次の積率を持つ確率変数とし

$$\boldsymbol{X} = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_d \end{pmatrix}$$

と書く. このとき確率ベクトル $\boldsymbol{X}$ の期待値を

$$\mathsf{E}[\boldsymbol{X}] := \begin{pmatrix} \mathsf{E}[X_1] \\ \mathsf{E}[X_2] \\ \vdots \\ \mathsf{E}[X_d] \end{pmatrix}$$

で定義する. $\boldsymbol{X}$ の共分散を

$$\mathsf{Var}[\boldsymbol{X}] := \mathsf{E}\big[(\boldsymbol{X}-\boldsymbol{\mu})(\boldsymbol{X}-\boldsymbol{\mu})^\top\big]$$

で定義する[10]. ただし $\boldsymbol{\mu} = (\mu_1,\, \mu_2,\, \ldots,\, \mu_d)^\top := \mathsf{E}[\boldsymbol{X}]$ である. これは

$$\mathsf{Var}[\boldsymbol{X}] = \begin{pmatrix} \mathsf{Var}[X_1] & \mathsf{Cov}[X_1,\, X_2] & \cdots & \mathsf{Cov}[X_1,\, X_d] \\ \mathsf{Cov}[X_2,\, X_1] & \mathsf{Var}[X_2] & \cdots & \mathsf{Cov}[X_2,\, X_d] \\ \vdots & \vdots & \ddots & \vdots \\ \mathsf{Cov}[X_d,\, X_1] & \mathsf{Cov}[X_d,\, X_2] & \cdots & \mathsf{Var}[X_d] \end{pmatrix}$$

である.

[9] これらより, $\mathsf{Pr}(X \le 1/2, Y \le 1/2) \ne \mathsf{Pr}(X \le 1/2)\mathsf{Pr}(Y \le 1/2)$ となり, $\mathsf{Cov}[X,\, Y] = 0$ だが, X と Y は従属であることがわかる.

[10] 確率ベクトルと同様に確率変数を成分とする行列を確率行列という. 確率行列の期待値はそれぞれの成分の期待値を取ったものを配置した行列と定義している.

注意 2.21. 定義から $\mathsf{Var}[\boldsymbol{X}]$ は半正定値対称行列となる. なぜならば, 期待値の線型性と補題 2.2(2) から, 任意の $\boldsymbol{a} \in \mathbb{R}^d$ に対して

$$\begin{aligned}\boldsymbol{a}^\top \mathsf{Var}[\boldsymbol{X}]\boldsymbol{a} &= \mathsf{E}\big[\boldsymbol{a}^\top(\boldsymbol{X}-\boldsymbol{\mu})(\boldsymbol{X}-\boldsymbol{\mu})^\top \boldsymbol{a}\big] \\ &= \mathsf{E}\big[\big\{\boldsymbol{a}^\top(\boldsymbol{X}-\boldsymbol{\mu})\big\}^2\big] \geq 0\end{aligned}$$

となる. 最後の不等号は補題 2.2(2) からわかる. よって $\mathsf{Var}[\boldsymbol{X}]$ は半正定値であることが示せた. □

補題 2.22. $\boldsymbol{X} = (X_1, X_2, \ldots, X_d)^\top$ は確率ベクトルで各成分は有限な 2 次の積率を持つとし

$$\mathsf{E}[\boldsymbol{X}] = \boldsymbol{\mu}, \quad \mathsf{Var}[\boldsymbol{X}] = \boldsymbol{\Sigma}$$

とする. ただし $\boldsymbol{\mu} \in \mathbb{R}^d$, $\boldsymbol{\Sigma}$ は $d \times d$ の半正定値行列 [11] である.
(1) 任意の定数ベクトル $\boldsymbol{a} \in \mathbb{R}^d$ に対して

$$\mathsf{E}[\boldsymbol{a}^\top \boldsymbol{X}] = \boldsymbol{a}^\top \boldsymbol{\mu}, \quad \mathsf{Var}[\boldsymbol{a}^\top \boldsymbol{X}] = \boldsymbol{a}^\top \boldsymbol{\Sigma} \boldsymbol{a}$$

となる.
(2) $k \in \mathbb{N}$ とする. 任意の定数の $k \times d$ 行列 $\mathbf{A}$ に対して

$$\mathsf{E}[\mathbf{A}\boldsymbol{X}] = \mathbf{A}\mathsf{E}[\boldsymbol{X}], \quad \mathsf{Var}[\mathbf{A}\boldsymbol{X}] = \mathbf{A}\mathsf{Var}[\boldsymbol{X}]\mathbf{A}^\top$$

となる.

Proof. (1) の証明: 期待値の線型性を用いて計算すればよい.
(2) の証明:

$$\mathbf{A} = \begin{bmatrix} \boldsymbol{a}_1^\top \\ \boldsymbol{a}_2^\top \\ \vdots \\ \boldsymbol{a}_k^\top \end{bmatrix}; \qquad \boldsymbol{a}_j \in \mathbb{R}^d \, (j = 1, 2, \ldots, k)$$

とおくと

$$\mathsf{E}\big[\mathbf{A}\boldsymbol{X}\big] = \mathsf{E}\begin{bmatrix} \boldsymbol{a}_1^\top \boldsymbol{X} \\ \boldsymbol{a}_2^\top \boldsymbol{X} \\ \vdots \\ \boldsymbol{a}_k^\top \boldsymbol{X} \end{bmatrix} = \begin{bmatrix} \mathsf{E}\big[\boldsymbol{a}_1^\top \boldsymbol{X}\big] \\ \mathsf{E}\big[\boldsymbol{a}_2^\top \boldsymbol{X}\big] \\ \vdots \\ \mathsf{E}\big[\boldsymbol{a}_k^\top \boldsymbol{X}\big] \end{bmatrix} = \begin{bmatrix} \boldsymbol{a}_1^\top \boldsymbol{\mu} \\ \boldsymbol{a}_2^\top \boldsymbol{\mu} \\ \vdots \\ \boldsymbol{a}_k^\top \boldsymbol{\mu} \end{bmatrix} = \mathbf{A}\boldsymbol{\mu}$$

[11] $d \times d$ の対称行列 $\mathbf{A}$ が半正定値であるとは, $\forall \boldsymbol{a} \in \mathbb{R}^d$ に対して $\boldsymbol{a}^\top \mathbf{A} \boldsymbol{a} \geq 0$ が成立するときをいう. また $d \times d$ の対称行列 $\mathbf{A}$ が正定値であるとは, $\forall \boldsymbol{a} \in \mathbb{R}^d \, (\boldsymbol{a} \neq \mathbf{0})$ に対して $\boldsymbol{a}^\top \mathbf{A} \boldsymbol{a} > 0$ が成立するときをいう.

がわかる. つぎに, 共分散の定義と上の結果から

$$
\begin{aligned}
\mathsf{Var}\big[\mathbf{A}\boldsymbol{X}\big] &= \mathsf{E}\big[\{\mathbf{A}\boldsymbol{X} - \mathsf{E}[\mathbf{A}\boldsymbol{X}]\}\{\mathbf{A}\boldsymbol{X} - \mathsf{E}[\mathbf{A}\boldsymbol{X}]\}^\top\big] \\
&= \mathsf{E}\big[\{\mathbf{A}\boldsymbol{X} - \mathbf{A}\mathsf{E}[\boldsymbol{X}]\}\{\mathbf{A}\boldsymbol{X} - \mathbf{A}\mathsf{E}[\boldsymbol{X}]\}^\top\big] \\
&= \mathsf{E}\big[\mathbf{A}\{\boldsymbol{X} - \mathsf{E}[\boldsymbol{X}]\}\{\boldsymbol{X} - \mathsf{E}[\boldsymbol{X}]\}^\top\mathbf{A}^\top\big] \\
&= \mathbf{A}\mathsf{E}\Big[\{\boldsymbol{X} - \mathsf{E}[\boldsymbol{X}]\}\{\boldsymbol{X} - \mathsf{E}[\boldsymbol{X}]\}^\top\Big]\mathbf{A}^\top \\
&= \mathbf{A}\mathsf{Var}[\boldsymbol{X}]\mathbf{A}^\top
\end{aligned}
$$

からわかる. □

問 2.4. $d = 2$ として, 補題 2.22 を確認せよ.

系 2.23. 確率ベクトル $\boldsymbol{X}$ は, 任意の定数ベクトル $\boldsymbol{a}(\neq \mathbf{0})$ に対して

$$\mathsf{Pr}(\boldsymbol{a}^\top\boldsymbol{X} = 0) = 0$$

をみたす. このとき, $\boldsymbol{\Sigma}$ は正定値である.

注意 2.24. $\boldsymbol{X}$ の共分散行列 $\boldsymbol{\Sigma}$ は, 補題 2.22(1) から半正定値であることがわかる. 系 2.23 の仮定は, 確率ベクトル $\boldsymbol{X}$ が d 次元空間より次元の低い空間に集中しないことを意味している. したがって, 確率ベクトル $\boldsymbol{X}$ の値を取る空間が退化していなければ, その共分散行列は正定値となることがわかる. □

2.4 条件付き期待値

定義 2.25. (1) X と Y を確率変数とし, 条件付き p.d.f.(または条件付き p.m.f.) を $\mathsf{p}^{X|Y}$(または $\mathsf{p}^{X|Y}$) とする. $Y = y$ を与えたときの X の条件付き期待値を

$$
\mathsf{E}[X|\,Y = y] = \begin{cases} \displaystyle\sum_x x\mathsf{p}^{X|Y}(x|\,y) & (\text{離散型の場合}) \\ \displaystyle\int_{-\infty}^{\infty} x\mathsf{p}^{X|Y}(x|\,y)\,\mathrm{d}x & (\text{連続型の場合}) \end{cases}
$$

で定義する. ただし考えている $Y = y$ で条件付き p.d.f. または p.m.f. は定義され, さらに $\mathsf{E}[|X|] < \infty$ とする.
(2) (Borel 可測) 関数 $g\colon \mathbb{R}^2 \to \mathbb{R}$ に対して, $g(X,\,Y)$ の条件付き期待値を

$$
\mathsf{E}[g(X,\,Y)|\,Y = y] = \begin{cases} \displaystyle\sum_x g(x,\,y)\mathsf{p}^{X|Y}(x|\,y) & (\text{離散型の場合}) \\ \displaystyle\int_{-\infty}^{\infty} g(x,\,y)\mathsf{p}^{X|Y}(x|\,y)\,\mathrm{d}x & (\text{連続型の場合}) \end{cases}
$$

で定義する. ただし, 考えている $Y = y$ での条件付き p.d.f. または p.m.f. は定義され, $\mathsf{E}[|g(X, Y)|] < \infty$ とする.

注意 2.26. 定義 2.1 と同様に関数 g を g^+ と g^- の部分に分けて条件付き期待値を定義をするべきであろうが, $\mathsf{E}[|g(X)|] < \infty$ の仮定のもとでは, $\mathsf{E}[g^+(X)|\,Y = y] < \infty$, $\mathsf{E}[g^-(X)|\,Y = y] < \infty$ が保証されるので, 上記のように簡単に定義をした.

注意 2.27. $\mathsf{E}[X]$ は定数であるが, $\mathsf{E}[X|\,Y = y]$ は一般に y の関数である. このことから

$$h(y) := \mathsf{E}[X|\,Y = y]$$

とおいたときに $h(y)$ に Y を代入したもの $h(Y)$ は確率変数 [12] になる. これを

$$\mathsf{E}[X|\,Y] := h(Y)$$

と記すことにする. したがって $\omega \in \Omega$ に対して, $y = Y(\omega)$ と書けば

$$\mathsf{E}[X|\,Y] : \Omega \ni \omega \mapsto \mathsf{E}[X|\,Y(\omega)] = \mathsf{E}[X|\,Y = y] \in \mathbb{R}$$

は可測となる.

測度論的確率論の教科書では, Radon-Nikodym の定理から条件付き期待値を定義する. これから条件付き p.d.f. を定義することになる. この点については [8, pp.181 − 182] を参照のこと. □

例 2.28. 連続型確率変数 Y は p.d.f.

$$\mathsf{p}^Y(y) = \begin{cases} 1 & (0 < y < 1) \\ 0 & (\text{その他の場合}) \end{cases}$$

を持つとする. $Y = y\,(0 < y < 1)$ を観測したとき

$$X|\,Y = y \sim \mathsf{Unif}(y, 1)$$

とする. すなわち $0 < y < 1$ のとき

$$\mathsf{p}^{X|Y}(x|\,y) = \begin{cases} \dfrac{1}{1-y} & (y < x < 1) \\ 0 & (\text{その他の場合}) \end{cases}$$

である. よって

$$\mathsf{E}[X|\,Y = y] = \int_y^1 x\mathsf{p}^{X|Y}(x|\,y)\,\mathrm{d}x = \frac{1}{1-y}\int_y^1 x\,\mathrm{d}x = \frac{1+y}{2}$$

[12] $h \mapsto h(y)$ の可測性は測度論の知識が必要となる. このことは定理 **??** から保証される.

となる．これより

$$\mathsf{E}[X|\,Y] = \frac{1+Y}{2}$$

となる．　□

定理 2.29. (1) 有限な期待値を持つ確率変数 X, Y と確率変数 Z に対して

$$\mathsf{E}[X+Y|\,Z] = \mathsf{E}[X|\,Z] + \mathsf{E}[Y|\,Z]$$

となる．
(2) 有限な期待値を持つ確率変数 X, Y に対して

$$\mathsf{E}[\mathsf{E}[Y|\,X]] = \mathsf{E}[Y], \quad \mathsf{E}[\mathsf{E}[X|\,Y]] = \mathsf{E}[X]$$

となる．
(3) 一般の (Borel 可測) 関数 $g : \mathbb{R}^2 \to \mathbb{R}$ を考える．$\mathsf{E}[|g(X, Y)|] < \infty$ のとき

$$\mathsf{E}[\mathsf{E}[g(X, Y)|\,Y]] = \mathsf{E}[g(X, Y)]$$

となる．
(4) $\mathsf{E}[XY|\,Y] = Y\mathsf{E}[X|\,Y]$ となる．

Proof. (1) は期待値の線型性よりわかる．連続型の場合について (2) の第 1 番目の等式を示す．他の場合もほとんど同じように証明できる．p を (X, Y) の同時 p.d.f. とする．このとき

$$\mathsf{p}(x, y) = \mathsf{p}^X(x)\mathsf{p}^{Y|X}(y|\,x)$$

となる．ただし p^X は X の周辺 p.d.f. で，$\mathsf{p}^{Y|X}$ は $X = x$ を与えたときの Y の条件付き p.d.f. である．$\mathsf{E}[Y|\,X = x] =: h(x)$ とおいたとき，$\mathsf{E}[Y|\,X] := h(X)$ と書いたことを思い出す．また，$S_X = \{x \in \mathbb{R};\, \mathsf{p}^X(x) >$

$0\}$ とおいたとき,

$$
\begin{aligned}
\mathsf{E}[\mathsf{E}[Y|\,X]] &= \mathsf{E}[h(X)] = \int_{-\infty}^{\infty} h(x)\mathsf{p}^X(x)\,\mathrm{d}x = \int_{S_X} h(x)\mathsf{p}^X(x)\,\mathrm{d}x \\
&= \int_{S_X} \mathsf{E}[Y|\,X=x]\mathsf{p}^X(x)\,\mathrm{d}x \\
&= \int_{S_X} \left\{\int_{-\infty}^{\infty} x\mathsf{p}^{Y|\,X}(y|\,x)\,\mathrm{d}y\right\} \mathsf{p}^X(x)\,\mathrm{d}x \\
&= \int_{-\infty}^{\infty}\int_{S_X} x\mathsf{p}^{Y|\,X}(y|\,x)\mathsf{p}^X(x)\,\mathrm{d}x\mathrm{d}y \\
&= \int_{-\infty}^{\infty}\left\{\int_{S_X} x\,\frac{\mathsf{p}(x,\,y)}{\mathsf{p}^X(x)}\,\mathsf{p}^X(x)\,\mathrm{d}x\right\}\mathrm{d}y \\
&= \int_{-\infty}^{\infty}\left\{\int_{S_X} x\,\mathsf{p}(x,\,y)\,\mathrm{d}x\right\}\mathrm{d}y \\
&= \int_{-\infty}^{\infty} x\left\{\int_{-\infty}^{\infty} \mathsf{p}(x,\,y)\,\mathrm{d}y\right\}\,\mathrm{d}x \\
&= \int_{-\infty}^{\infty} x\mathsf{p}^X(x)\,\mathrm{d}x \\
&= \mathsf{E}[X]
\end{aligned}
$$

となる [13]. 積分の順序交換は $g(X,\,Y)$ が有限の期待値を持つことから保証されること [14] が知られている. (3) と (4) も同様に証明できる. □

問 2.5. 定理 2.29(3)(4) の証明を書け.

例 2.30. (例 2.28 の続き)

$$
\mathsf{E}[X] = \mathsf{E}\left[\frac{1+Y}{2}\right] = \frac{3}{4}
$$

となる. 一方 $0<y<x<1$ に対して

$$
\mathsf{p}(x,\,y) = \mathsf{p}^{X|\,Y}(x|\,y)\mathsf{p}^Y(y) = \frac{1}{1-y}
$$

となる. よって

$$
\begin{aligned}
\mathsf{E}[X] &= \int_0^1 \left\{\int_y^1 x\,\frac{1}{1-y}\,\mathrm{d}x\right\}\,\mathrm{d}y \\
&= \int_0^1 \frac{1}{1-y}\left[\frac{x^2}{2}\right]_y^1\,\mathrm{d}y \\
&= \int_0^1 \frac{1}{1-y}\,\frac{1-y^2}{2}\,\mathrm{d}y = \left[\frac{(1+y)^2}{4}\right]_0^1 = \frac{3}{4}
\end{aligned}
$$

[13] h は $S_Y := \{y \in \mathbb{R};\ \mathsf{p}^Y(y) > 0\}$ においてのみ定義されるが, 形式的に $h(y) = 0\,(y \in S_Y)$ と考えて上の式を理解すればよい.

[14] Fubini の定理からわかる.

となる. □

定義 2.31. X は有限の 2 次の積率を持つとする. $Y = y$ を与えたときの X の条件付き p.d.f. $\mathsf{p}^{X|Y}$(p.m.f. $\mathsf{p}^{X|Y}$) が定義できる y を考える. このとき, $Y = y$ を与えたときの条件付き分散を

$$\mathsf{Var}[X|\,Y=y] = \begin{cases} \displaystyle\sum_x \{x-\mu(y)\}^2 \mathsf{p}^{X|Y}(x|\,y) & \text{(離散型の場合)} \\ \displaystyle\int_{-\infty}^{\infty} \{x-\mu(y)\}^2 \mathsf{p}^{X|Y}(x|\,y)\,\mathrm{d}x & \text{(連続型の場合)} \end{cases}$$

で定義する. ただし $\mu(y) = \mathsf{E}[X|\,Y=y]$ である. これは

$$\mathsf{Var}[X|\,Y] = \mathsf{E}[X^2|\,Y] - \{\mathsf{E}[X|\,Y]\}^2$$

とも書ける.

定理 2.32. X, Y を確率変数とし $\mathsf{E}[X^2] < \infty$ とする. このとき

$$\mathsf{Var}[X] = \mathsf{E}[\mathsf{Var}[X|\,Y]] + \mathsf{Var}[\mathsf{E}[X|\,Y]]$$

が成立する. ただし $h(y) := \mathsf{Var}[X|\,Y=y]$ としたとき $\mathsf{Var}[X|\,Y] := h(Y)$ と定義した.

Proof. まず

$$\begin{aligned} \mathsf{Var}[X] &= \mathsf{E}\big[\{X-\mathsf{E}[X]\}^2\big] \\ &= \mathsf{E}\Big[\big\{X-\mathsf{E}[X|\,Y]+\mathsf{E}[X|\,Y]-\mathsf{E}[X]\big\}^2\Big] \\ &= \mathsf{E}\big[\{X-\mathsf{E}[X|\,Y]\}^2\big] + \mathsf{E}\big[\{\mathsf{E}[X|\,Y]-\mathsf{E}[X]\}^2\big] \\ &\quad + 2\mathsf{E}\big[\{X-\mathsf{E}[X|\,Y]\}\{\mathsf{E}[X|\,Y]-\mathsf{E}[X]\}\big] \end{aligned} \tag{2.3}$$

と書き直す. しかし, (2.3) の最右辺の各項は以下のように評価できる.

$$
\begin{aligned}
\mathsf{E}\big[\{X-\mathsf{E}[X|Y]\}^2\big] &= \mathsf{E}\big[X^2-2X\mathsf{E}[X|Y]+\{\mathsf{E}[X|Y]\}^2\big] \\
&= \mathsf{E}\Bigg[\mathsf{E}\big[X^2-2X\mathsf{E}[X|Y]+\{\mathsf{E}[X|Y]\}^2\big]\Bigg|\, Y\Bigg] \\
&\quad (\because \text{定理 } 2.29(3)) \\
&= \mathsf{E}\Bigg[\mathsf{E}\big[X^2\big|\,Y\big]-2\mathsf{E}[X|Y]\,\mathsf{E}[X|Y]+\{\mathsf{E}[X|Y]\}^2\Bigg] \\
&= \mathsf{E}\Bigg[\mathsf{E}[X^2|Y]-\{\mathsf{E}[X|Y]\}^2\Bigg] \\
&= \mathsf{E}\big[\mathsf{Var}[X|Y]\big], \\
\mathsf{E}\Bigg[\{\mathsf{E}[X|Y]-\mathsf{E}[X]\}^2\Bigg] &= \mathsf{E}\Bigg[\Big\{\mathsf{E}[X|Y]-\underbrace{\mathsf{E}[\mathsf{E}[X|Y]]}_{=\mathsf{E}[X]}\Big\}^2\Bigg] \\
&\quad (\because \text{定理 } 2.29(3)) \\
&= \mathsf{Var}\Bigg[\mathsf{E}[X|Y]\Bigg], \\
\mathsf{E}\big[\{X-\mathsf{E}[X|Y]\}\{\mathsf{E}[X|Y]-\mathsf{E}[X]\}\big] &= \mathsf{E}\Bigg[\mathsf{E}\big[\{X-\mathsf{E}[X|Y]\}\{\mathsf{E}[X|Y]-\mathsf{E}[X]\}\Bigg|\,Y\Bigg] \\
&\quad (\because \text{定理 } 2.29(3)) \\
&= \mathsf{E}\Bigg[\{\mathsf{E}[X|Y]-\mathsf{E}[X]\}\,\underbrace{\mathsf{E}\big[X-\mathsf{E}[X|Y]\big|Y\big]}_{=0}\Bigg] \\
&= 0
\end{aligned}
$$

となる. 最後から 2 番目の等号は定理 2.29(3) を用いた. これらの結果を (2.3) の最右辺の各項に代入すれば, 定理は証明される. □

2.5 積率母関数

定義 2.33. X を確率変数とし, ある $t_0>0$ が存在して, $\mathsf{E}[e^{tX}]<\infty$ $(\forall|t|<t_0)$ とする. このとき, X の**積率母関数** (Moment Generating Function (m.g.f.)) を

$$\mathsf{m}^X(t) := \mathsf{E}[e^{tX}] \quad (-t_0<t<t_0)$$

と定義する.

注意 2.34. 確率変数 X の積率母関数 $\mathsf{m}^X(t)$ が存在するとき, 期待値と微

分の記号の入れ替えが保証されること [15] が知られている．このことから

$$\dot{\mathsf{m}}^X(0) = \dot{\mathsf{m}}^X(t)\Big|_{t=0} = \left[\frac{\mathrm{d}}{\mathrm{d}t} m^X(t)\right]\Bigg|_{t=0} = \left[\frac{\mathrm{d}}{\mathrm{d}t}\mathsf{E}\big[e^{tX}\big]\right]\Bigg|_{t=0} = \mathsf{E}\left[\frac{\mathrm{d}}{\mathrm{d}t} e^{tX}\right]\Bigg|_{t=0}$$
$$= \mathsf{E}\big[Xe^{tX}\big]\Big|_{t=0} = \mathsf{E}[X]$$

となる．この議論を繰り返せば $k = 2, 3, \ldots$ に対して

$$\{\mathsf{m}^X\}^{(k)}(0) = \mathsf{E}\big[X^k\big]$$

がわかる． □

例 2.35. $X \sim \mathsf{Exp}(1)$ とする．$t < 1$ に対して

$$\mathsf{m}^X(t) = \mathsf{E}\big[e^{tX}\big] = \int_0^\infty e^{tx} e^{-x}\,\mathrm{d}x = \int_0^\infty e^{-(1-t)x}\,\mathrm{d}x = \frac{1}{1-t}$$

となる．$t \geq 1$ のときは, e^{tX} の期待値は発散する．したがって

$$\mathsf{m}^X(t) = \frac{1}{1-t} \qquad (t < 1)$$

となる．簡単な計算から

$$\dot{\mathsf{m}}^X(0) = 1, \quad \ddot{\mathsf{m}}^X(0) = 2$$

なので

$$\mathsf{E}[X] = 1, \quad \mathsf{E}[X^2] = 2, \quad \mathsf{Var}[X] = \mathsf{E}[X^2] - \{\mathsf{E}[X]\}^2 = 1$$

となる． □

補題 2.36. (1) $a, b \in \mathbb{R}\,(a \neq 0)$ とする．$Y = aX + b$ としたとき

$$\mathsf{m}^Y(t) = e^{tb}\mathsf{m}^X(at)$$

となる．
(2) $X_1, X_2, \ldots, X_d$ は独立とし $Y = \sum_{j=1}^d X_j$ とする．このとき

$$\mathsf{m}^Y(t) = \prod_{j=1}^d \mathsf{m}^{X_j}(t)$$

となる．

[15]たとえば, [39, pp.75-76] を参照のこと．

Proof. (1) は指数関数と期待値の性質よりわかる. (2) は $X_j\,(j=1,\,2,\,\ldots,\,d)$ の独立性と指数関数の性質に注意して, 定理 2.12(3) を適用するとわかる. □

定理 2.37. X と Y を確率変数とする. ある数 $t_0>0$ が存在して

$$\mathsf{m}^X(t)=\mathsf{m}^Y(t)\quad(|t|<t_0)$$

ならば

$$X\overset{d}{=}Y$$

となる. ただし X と Y の c.d.f. を F^X と F^Y としたとき

$$X\overset{d}{=}Y\ \Leftrightarrow\ \mathsf{F}^X(x)=\mathsf{F}^Y(x)\,(\forall x\in\mathbb{R})$$

である.

Proof. これは信じることにする. □

注意 2.38. 定理 2.37 の証明は, 積率母関数が存在する範囲に対する $\mathbb{C}$ の帯領域に解析接続し, それに対して, Fourier 逆変換の公式を適応して証明するのが標準的であろう.

例 2.39. $n_1,\,n_2\in\mathbb{N},0<p<1$ とし $X_1\sim\mathsf{Bino}(n_1,\,p)$, $X_2\sim\mathsf{Bino}(n_2,\,p)$ は独立とする. $Y=X_1+X_2$ としたとき

$$\mathsf{m}^Y(t)=\mathsf{m}^{X_1}(t)\mathsf{m}^{X_2}(t)=\left(pe^t+q\right)^{n_1}\left(pe^t+q\right)^{n_2}=(pe^t+q)^{n_1+n_2}$$

となる. ただし $q=1-p$ である. よって $Y\sim\mathsf{Bino}(n_1+n_2,\,p)$ となる. □

例 2.40. $\lambda_1,\,\lambda_2>0$ とし, $X_1\sim\mathsf{Po}(\lambda_1)$, $X_2\sim\mathsf{Po}(\lambda_2)$ は独立とする. $Y=X_1+X_2$ としたとき

$$\mathsf{m}^Y(t)=\mathsf{m}^{X_1}(t)\mathsf{m}^{X_2}(t)=e^{\lambda_1(e^t-1)}e^{\lambda_2(e^t-1)}=e^{(\lambda_1+\lambda_2)(e^t-1)}$$

となる. したがって, $Y\sim\mathsf{Po}(\lambda_1+\lambda_2)$ がわかる. □

2.6 章末注釈と参考文献

この章は [36] を参考にした.

2.7 演習問題

演習問題 2.1. 確率空間 $(\Omega, \mathcal{A}, \mathsf{Pr})$ 上の確率変数 X は

$$\mathsf{Pr}(X = c) = 1; \qquad c \text{ は定数}$$

をみたすとする. このとき, X の平均 $\mathsf{E}[X]$ と分散 $\mathsf{Var}[X]$ を求めよ.

演習問題 2.2. (1) $X \sim \mathsf{Ber}(\theta)\,(0 < \theta < 1)$ のとき

$$\mathsf{E}[X] = \theta, \quad \mathsf{Var}[X] = \theta(1 - \theta)$$

を示せ.
(2) $X_1, X_2, \ldots, X_n \sim \text{i.i.d.}\mathsf{Ber}(\theta)\,(0 < \theta < 1)$ とする. $S = \sum_{j=1}^{n} X_j$ とおいたとき, $S \sim \mathsf{Bino}(n, \theta)$ となることを示せ.
(3) $S \sim \mathsf{Bino}(n, \theta)\,(n \in \mathbb{N}, 0 < \theta < 1)$ のとき

$$\mathsf{E}[S] = n\theta, \quad \mathsf{Var}[S] = n\theta(1 - \theta)$$

を示せ.
(4) $X \sim \mathsf{Po}(\theta)\,(\theta > 0)$ とする. このとき, X の平均 $\mathsf{E}[X]$ と分散 $\mathsf{Var}[X]$ を求めよ.
(5) $U \sim \mathsf{U}(0, 1)$ のとき

$$\mathsf{E}[X] = \frac{1}{2}, \quad \mathsf{Var}[X] = \frac{1}{12}$$

を示せ.
(6) $X \sim \mathsf{Ex}(\theta)\,(\theta > 0)$ とする. このとき

$$\mathsf{E}[X] = \frac{1}{\theta}, \quad \mathsf{Var}[X] = \frac{1}{\theta^2}$$

を示せ.
(6) $X \sim \mathsf{N}(\mu, \sigma^2)\,(\mu \in \mathbb{R}, \sigma > 0)$ とする. このとき

$$\mathsf{E}[X] = \mu, \quad \mathsf{Var}[X] = \sigma^2$$

を示せ.
(7) $X \sim \chi^2_n$ のとき

$$\mathsf{E}[X] = n, \quad \mathsf{E}[X] = 2n$$

を示せ.

演習問題 2.3. 確率空間 $(\Omega, \mathcal{A}, \mathsf{Pr})$ 上の確率変数 X は $\mathsf{Var}[X] = \sigma^2\,(0 < \sigma < \infty)$ とする. c を定数としたとき, $\mathsf{Var}[X + c]$ と $\mathsf{Var}[cX]$ を σ と c を用いて表せ.

演習問題 2.4. (1) $X \sim \mathsf{Ga}(\alpha, \beta)$ とする. X の k 次の積率は以下で与えられることを示せ.

$$\mathsf{E}[X^k] = \frac{(a+k-1)(a+k-2)\cdots\alpha}{\beta^k}.$$

(2) 上の問いの結果から

$$\mathsf{E}[X] = \frac{\alpha}{\beta}, \quad \mathsf{E}[X^2] = \frac{(\alpha+1)\alpha}{\beta^2}, \quad \mathsf{Var}[X] = \frac{\alpha}{\beta^2}$$

を示せ.
(3) X の積率母関数は以下で与えられることを示せ.

$$\mathsf{m}^X(t) = \left(\frac{1}{1-\beta t}\right)^\alpha \quad (t < 1/\beta).$$

演習問題 2.5. X を確率空間 $(\Omega, \mathcal{A}, \mathsf{Pr})$ 上の連続型確率変数とし, X の p.d.f. を

$$\mathsf{p}(x) = \begin{cases} 2x & (0 < x < 1) \\ 0 & (\text{その他の場合}) \end{cases}$$

とする. さらに, 離散型確率変数 Y を

$$Y := \mathbb{1}_{(1/2,\,\infty)}(X) := \begin{cases} 1 & \left(X > \frac{1}{2}\right) \\ 0 & \left(X \le \frac{1}{2}\right) \end{cases}$$

で定める. このとき, 以下の問に答えよ.
(1) 確率変数 X の平均 $\mathsf{E}[X]$ と分散 $\mathsf{Var}[X]$ を求めよ.
(2) 確率変数 Y の平均 $\mathsf{E}[Y]$ と分散 $\mathsf{Var}[Y]$ を求めよ.
(3) 確率変数 X と Y の共分散 $\mathsf{Cov}[X, Y]$ を求めよ.

演習問題 2.6. X_1, X_2 を確率空間 $(\Omega, \mathcal{A}, \mathsf{Pr})$ 上の連続型確率変数とし,

$$X_1 \sim \mathsf{N}(x_0+\mu, \sigma^2), \qquad X_2 |\, X_1 = x_1 \sim \mathsf{N}(x_1+\mu, \sigma^2)$$

とする. ただし, $x_0, x_1, \mu \in \mathbb{R}$ である. このとき, 以下の問いに答えよ. ただし, 正規分布 $\mathsf{N}(\mu, \sigma^2)$ に従う確率変数 Z の期待値と分散が $\mathsf{E}[Z] = \mu$, $\mathsf{Var}[Z] = \sigma^2$ となることは証明なしで用いてよい. また, 期待値, 分散, 共分散, 条件付き期待値に係る資料に書いてある性質も証明なしで用いてよい. なお, どの性質を用いたかは明示すること.
(1) 確率変数 $X_2 - X_1$ の期待値 $\mathsf{E}[X_2 - X_1]$ を求めよ.
(2) X_2 の分散 $\mathsf{Var}[X_2]$ を求めよ.
(3) X_1 と X_2 の共分散 $\mathsf{Cov}[X_1, X_2]$ を求めよ.
(4) $X_2 - X_1$ の分散 $\mathsf{Var}[X_2 - X_1]$ を $\mathsf{Var}[X_1]$, $\mathsf{Var}[X_2]$, $\mathsf{Cov}[X_1, X_2]$ で表現せよ.
(5) $X_2 - X_1$ の分散 $\mathsf{Var}[X_2 - X_1]$ を求めよ.

演習問題 2.7. $(\Omega,\, \mathcal{A},\, \mathsf{Pr})$ を確率空間とし, X をこの空間上で定義された確率変数とする. X は開区間 $(0,\, 1)$ 上の一様分布に従うとし

$$\boldsymbol{X} = \begin{bmatrix} X \\ X^2 \end{bmatrix}$$

とおく. このとき, 以下の問いに答えよ.
(1) 確率ベクトル $\boldsymbol{X}$ の期待値 $\mathsf{E}[\boldsymbol{X}]$ を求めよ.
(2)

$$\mathsf{Var}[\boldsymbol{X}] = \mathsf{E}[\boldsymbol{X}\boldsymbol{X}^\top] - \mathsf{E}[\boldsymbol{X}]\big\{\mathsf{E}[\boldsymbol{X}]\big\}^\top$$

を示せ.
(3) 確率ベクトル $\boldsymbol{X}$ の分散共分散行列 $\mathsf{Var}[\boldsymbol{X}]$ を求めよ.
(4) $\det\big[\mathsf{Var}[\boldsymbol{X}]\big]$ を求めよ.

演習問題 2.8. 確率空間 $(\Omega,\, \mathcal{A},\, \mathsf{Pr})$ 上で定義された連続型確率ベクトル $(X,\, Y)$ は同時 p.d.f.

$$\mathsf{p}^{(X,\,Y)}(x,\, y) = \begin{cases} \dfrac{3}{8} \times y & (0 \leq x \leq y \leq 2) \\ 0 & (\text{その他の場合}) \end{cases}$$

を持つとする.
(1) X の周辺 p.d.f. $\mathsf{p}^X(x)$ を求め,

$$\int_{-\infty}^{\infty} \mathsf{p}^X(x)\,\mathrm{d}x = 1$$

を確認せよ.
(2) Y の周辺 p.d.f. $\mathsf{p}^Y(y)$ を求め,

$$\int_{-\infty}^{\infty} \mathsf{p}^Y(y)\,\mathrm{d}y = 1$$

を確認せよ.
(3) X と Y の期待値 $\mathsf{E}[X]$ と $\mathsf{E}[Y]$ を求めよ.
(4) XY の期待値 $\mathsf{E}[XY]$ を求め, X と Y の共分散 $\mathsf{Cov}[X,\, Y]$ を求めよ.
(5) $\mathsf{p}^Y(y) > 0$ なる y に対して, $Y = y$ を与えたときの X の条件付き p.d.f. $\mathsf{p}^{X|Y}(x|\,y)$ を求めよ. $\mathsf{p}^{X|Y}(x|\,y) > 0$ となる x の範囲を書いたうえで,

$$\int_{-\infty}^{\infty} \mathsf{p}^{X|Y}(x|\,y)\,\mathrm{d}x = 1$$

を確認せよ.

演習問題 2.9. 確率空間 $(\Omega,\, \mathcal{A},\, \mathsf{Pr})$ 上で定義された連続型確率ベクトル $(X,\, Y)$ は同時 p.d.f.

$$\mathsf{p}^{(X,Y)}(x,\, y) = \begin{cases} e^{-(x+y)} & (0 < x < \infty,\, 0 < y < \infty) \\ 0 & (\text{その他の場合}) \end{cases}$$

を持つ.

$$U = 2X, \qquad V = X + Y$$

とおく.
(1) 確率ベクトル $(U,\, V)$ の同時 p.d.f. $\mathsf{p}^{(U,V)}(u,\, v)$ を求めよ. さらに

$$T = \{(u,\, v) \in \mathbb{R}^2 :\, \mathsf{p}^{(U,V)}(u,\, v) > 0\}$$

を図示せよ.
(2) U の周辺 p.d.f. $\mathsf{p}^U(u)$ を求めよ.

演習問題 2.10. $\lambda > 0$ とし, W は確率空間 $(\Omega,\, \mathcal{A},\, \mathsf{Pr})$ 上で定義された確率変数で, 母数 λ の指数分布 $\mathsf{Exp}(\lambda)$ に従うとする. すなわち, W は連続型確率変数で, p.d.f.

$$\mathsf{p}^W(w) = \begin{cases} \lambda e^{-\lambda w} & (w > 0) \\ 0 & (\text{その他の場合}) \end{cases}$$

をもつ. 同じ確率空間 $(\Omega,\, \mathcal{A},\, \mathsf{Pr})$ 上で定義された確率変数 X と Y を

$$X = \lfloor W \rfloor, \quad Y = W - X$$

で定める. ただし, $\lfloor x \rfloor$ は実数 x 以下の最大の整数とする. たとえば, $\lfloor 1.5 \rfloor = 1$ となる. このとき, 以下の問いに答えよ.
(1) 非負の整数 x に対して

$$\mathsf{Pr}(X = x)$$

を求めることで X の p.m.f. p^X を求めよ.
(2) $\displaystyle\sum_{x=0}^{\infty} \mathsf{p}^X(x)$ を計算せよ.
(3) x を非負の整数とし, $0 < y < 1$ とする. $X = x$ を与えたときの $\{Y \leq y\}$ の条件付き確率

$$\mathsf{Pr}(Y \leq y|\, X = x)$$

を求めよ.
(4) x を非負の整数とする. $y \in \mathbb{R}$ に対して, $\mathsf{F}^{Y|X}(y|\, x) = \mathsf{Pr}(Y \leq y|\, X = x)$ と定める. $\mathsf{F}^{Y|X}(y|\, x)$ を実数上で定義された関数として書き表せ.

(5) x を非負の整数とする. $X = x$ を与えたときの Y の条件付き p.d.f. $\mathsf{p}^{Y|X}(y|\,x)$ を求めよ.
(6) 期待値 $\mathsf{E}[Y]$ を求めよ.

$\boxed{\text{補足}}$ $\mathsf{Pr}(0 < Y < 1) = 1$ から $0 < \mathsf{E}[Y] < 1$ となる. 得られた結果について, 不等式 $0 < \mathsf{E}[Y] < 1$ となっているかを確認するとよい.

演習問題 2.11. X を確率空間 $(\Omega, \mathcal{A}, \mathsf{Pr})$ 上の連続型確率変数とし, X の p.d.f. を

$$\mathsf{p}(x) = \begin{cases} 2x & (0 < x < 1) \\ 0 & (\text{その他の場合}) \end{cases}$$

とする. さらに, 離散型確率変数 Y を

$$Y := \mathbb{1}_{(1/2,\,\infty)}(X) := \begin{cases} 1 & \left(X > \frac{1}{2}\right) \\ 0 & \left(X \leq \frac{1}{2}\right) \end{cases}$$

で定める. このとき, 以下の問に答えよ.
(1) 確率変数 X の平均 $\mathsf{E}[X]$ と分散 $\mathsf{Var}[X]$ を求めよ.
(2) 確率変数 Y の平均 $\mathsf{E}[Y]$ と分散 $\mathsf{Var}[Y]$ を求めよ.
(3) 確率変数 X と Y の共分散 $\mathsf{Cov}[X, Y]$ を求めよ.

演習問題 2.12. (U, Y) の同時分布を以下のように定める. $U \sim \mathsf{Unif}(0, 1)$ とし, $U = u\,(0 < u < 1)$ が与えられたときの Y の条件付き分布は $\mathsf{N}(\theta, \sigma^2)$ とする. ただし,

$$\theta = 2 - 6u, \quad \sigma^2 = 25$$

とする.
(1) U の期待値 $\mathsf{E}[U]$ を計算せよ.
(2) U の分散 $\mathsf{Var}[U]$ を計算せよ.
(3) $U = u\,(0 < u < 1)$ を与えたときの Y の条件付き期待値 $\mathsf{E}[Y|\,U = u]$ を答えよ.
(4) Y の期待値 $\mathsf{E}[Y]$ を計算せよ.
(5) Y の分散 $\mathsf{Var}[Y]$ を求めよ.
(6) $U = u\,(0 < u < 1)$ を与えたときの条件付き分散 $\mathsf{Var}[Y|\,U = u]$ を答えよ.
(7) $\mathsf{E}\big[\mathsf{Var}[Y|\,U]\big] + \mathsf{Var}\big[\mathsf{E}[Y|\,U]\big]$ を計算せよ.
(8) 連続関数 $g\colon [0, 1] \to \mathbb{R}$ に対して

$$\mathsf{E}\Big[\big(Y - g(U)\big)^2\Big] \geq 25$$

となること示せ.

演習問題 2.13. $X \sim \mathsf{N}(0, 1)$ とする. $\mathsf{Var}[g(X)] < \infty$ なる 絶対連続関数 $g : \mathbb{R} \longrightarrow \mathbb{R}$ に対して

$$\mathsf{Var}[g(X)] \leq \mathsf{E}\big[\{\overset{\bullet}{g}(X)\}^2\big]$$

が成立することを示せ.

演習問題 2.14. 確率変数 X と Y は独立同一に共通の分布関数 F と p.d.f. p を持つ分布に従うとする.
(1) $V = \max X, Y$ は分布関数 $\mathsf{F}^V(v) = \{\mathsf{F}(v)\}^2\ (v \in \mathbb{R})$ と p.d.f. $\mathsf{p}^V(v) = 2\mathsf{p}(v)\mathsf{F}(v)$ を持つことを示せ.
(2) $U = \min\{X, Y\}$ の分布関数 F^U と p.d.f. p^U を求めよ.
(3) X と Y は独立同一に $[0, 1]$ 上の一様分布に従うとき, $\mathsf{E}[U]$ と $\mathsf{Cov}[U, V]$ を求めよ.
(4) X と Y は独立同一に指数分布

$$\mathsf{p}(x) = \begin{cases} e^{-x} & (x > 0) \\ 0 & (\text{その他の場合}) \end{cases}$$

に従うとき, U の分布を求めよ. さらに, $\mathsf{E}[V]$ と $\mathsf{Var}[V]$ を求めよ.