定義：

$g(X)$ の期待値 $\mathrm{E}(g(X))$：

$$\mathrm{E}(g(X)) = \begin{cases} \displaystyle\sum_i g(x_i) p_i = \sum_i g(x_i) f(x_i), \\ \qquad\qquad\qquad\qquad \text{離散型確率変数} \\ \displaystyle\int_{-\infty}^{\infty} g(x) f(x) \mathrm{d}x, \\ \qquad\qquad\qquad\qquad \text{連続型確率変数} \end{cases}$$

1. 確率変数 X の平均 $\mathrm{E}(X)$

$\Longrightarrow$ X の期待値, $g(X) = X$

$$
\mathrm{E}(\boldsymbol{X}) = \begin{cases} \displaystyle\sum_i \boldsymbol{x_i f(x_i)}, & \text{離散型確率変数} \\ \displaystyle\int_{-\infty}^{\infty} \boldsymbol{xf(x)}\mathrm{d}\boldsymbol{x}, & \text{連続型確率変数} \end{cases}
$$

$$
= \boldsymbol{\mu}, \quad (\text{または，} \boldsymbol{\mu_x})
$$

2. 確率変数 $\boldsymbol{X}$ の分散 $\mathrm{V}(\boldsymbol{X})$

$\Longrightarrow (\boldsymbol{X} - \boldsymbol{\mu})^2$ の期待値, $\boldsymbol{g(X)} = (\boldsymbol{X} - \boldsymbol{\mu})^2$

$$
\mathrm{V}(\boldsymbol{X}) = \mathrm{E}((\boldsymbol{X} - \boldsymbol{\mu})^2)
$$

$$
= \begin{cases} \displaystyle\sum_i (x_i - \mu)^2 f(x_i), \\ \qquad\qquad\qquad\qquad\text{離散型確率変数} \\ \displaystyle\int_{-\infty}^{\infty} (x - \mu)^2 f(x)\mathrm{d}x, \\ \qquad\qquad\qquad\qquad\text{連続型確率変数} \end{cases}
$$

$$
= \sigma^2, \quad (\text{または，} \sigma_x^2)
$$

確率変数 X の分散 $\mathrm{V}(X)$

$\Longrightarrow$ X の確率分布の確率関数 (離散型の場合)，または，確率密度関数 (連続型の場合) の範囲が広ければ，$\mathrm{V}(X)$ は大きい。

いくつかの公式：

1. a, b を定数とする。

定理： $\mathrm{E}(aX+b)=a\mathrm{E}(X)+b$

証明：

X が離散型確率変数の場合，

$$
\begin{aligned}
\mathrm{E}(aX+b) &= \sum_i (ax_i+b)f(x_i) \\
&= a\sum_i x_i f(x_i) + b\sum_i f(x_i) \\
&= a\mathrm{E}(X)+b
\end{aligned}
$$

途中で，$\displaystyle\sum_i f(x_i)=1$ に注意

X が連続型確率変数の場合，

$$
\mathrm{E}(aX+b)=\int_{-\infty}^{\infty}(ax+b)f(x)\mathrm{d}x
$$

$$
\begin{aligned}
&= \boldsymbol{a}\int_{-\infty}^{\infty} \boldsymbol{x}\boldsymbol{f}(\boldsymbol{x})\mathrm{d}\boldsymbol{x} + \boldsymbol{b}\int_{-\infty}^{\infty} \boldsymbol{f}(\boldsymbol{x})\mathrm{d}\boldsymbol{x} \\
&= \boldsymbol{a}\mathrm{E}(\boldsymbol{X}) + \boldsymbol{b}
\end{aligned}
$$

途中で，$\displaystyle\int_{-\infty}^{\infty} \boldsymbol{f}(\boldsymbol{x})\mathrm{d}\boldsymbol{x} = \mathbf{1}$ に注意

2. 定理： $\mathrm{V}(\boldsymbol{X}) = \mathrm{E}(\boldsymbol{X}^{2}) - \boldsymbol{\mu}^{2}$

証明：

$$
\begin{aligned}
\mathrm{V}(\boldsymbol{X}) &= \mathrm{E}((\boldsymbol{X} - \boldsymbol{\mu})^{2}) \\
&= \mathrm{E}(\boldsymbol{X}^{2} - 2\boldsymbol{\mu}\boldsymbol{X} - \boldsymbol{\mu}^{2}) \\
&= \mathrm{E}(\boldsymbol{X}^{2}) - 2\boldsymbol{\mu}\mathrm{E}(\boldsymbol{X}) + \boldsymbol{\mu}^{2} \\
&= \mathrm{E}(\boldsymbol{X}^{2}) - \boldsymbol{\mu}^{2}
\end{aligned}
$$

途中で，$\boldsymbol{\mu} = \mathrm{E}(\boldsymbol{X})$ に注意

3. a, b を定数とする。

定理： $\mathrm{V}(aX+b)=a^2\mathrm{V}(X)$

証明：

$\mathrm{E}(aX+b)=a\mu+b$ に注意して，

$$\begin{aligned}
\mathrm{V}(aX+b) &= \mathrm{E}(((aX+b)-\mathrm{E}(aX+b))^2) \\
&= \mathrm{E}((aX-a\mu)^2) \\
&= \mathrm{E}(a^2(X-\mu)^2) \\
&= a^2\mathrm{E}((X-\mu)^2) \\
&= a^2\mathrm{V}(X)
\end{aligned}$$

を得る。

4. 定理： 確率変数 X について，$\mathrm{E}(X)=\mu$, $\mathrm{V}(X)=\sigma^2$ とする。

$Z = \dfrac{X - \mu}{\sigma}$ を定義する。

このとき，$\mathrm{E}(Z) = 0$, $\mathrm{V}(Z) = 1$ となる。

証明：

$$\mathrm{E}(Z) = \mathrm{E}\left(\frac{X - \mu}{\sigma}\right) = \frac{\mathrm{E}(X) - \mu}{\sigma} = 0$$

$$\mathrm{V}(Z) = \mathrm{V}\left(\frac{1}{\sigma}X - \frac{\mu}{\sigma}\right) = \frac{1}{\sigma^2}\mathrm{V}(X) = 1$$

2 変数： 確率変数 X, Y のある関数： $g(X, Y)$

定義：

$\boldsymbol{g(X, Y)}$ の期待値 $\mathbf{E}(\boldsymbol{g(X, Y)})$：

$$\mathrm{E}(g(X,Y)) = \begin{cases} \displaystyle\sum_i \sum_j g(x_i, y_j) f(x_i, y_j), & \text{離散型確率変数} \\ \displaystyle\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f(x, y) \mathrm{d}x \mathrm{d}y, & \text{連続型確率変数} \end{cases}$$

1. 確率変数 X の平均 $\mathrm{E}(X)$

$\Longrightarrow$ X の期待値, $g(X, Y) = X$

$$\mathrm{E}(\boldsymbol{X}) = \begin{cases} \displaystyle\sum_i \sum_j \boldsymbol{x_i f(x_i, y_j)}, & \\ & \text{離散型確率変数} \\ \displaystyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \boldsymbol{xf(x,y)}\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{y}, & \\ & \text{連続型確率変数} \end{cases}$$

$$= \boldsymbol{\mu_x}$$

2. 確率変数 $\boldsymbol{X}$ の分散 $\mathrm{V}(\boldsymbol{X})$

$\Longrightarrow (\boldsymbol{X} - \boldsymbol{\mu_x})^2$ の期待値, $\boldsymbol{g(X, Y)} = (\boldsymbol{X} - \boldsymbol{\mu_x})^2$

$$\mathrm{V}(\boldsymbol{X}) = \mathrm{E}((\boldsymbol{X} - \boldsymbol{\mu_x})^2)$$

$$= \begin{cases} \displaystyle\sum_i \sum_j (x_i - \mu_x)^2 f(x_i, y_j), \\ \hfill \text{離散型確率変数} \\ \displaystyle\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mu_x)^2 f(x, y) \mathrm{d}x \mathrm{d}y, \\ \hfill \text{連続型確率変数} \end{cases}$$

$$= \sigma_x^2$$

3. 確率変数 X, Y の共分散 $\mathrm{Cov}(X, Y)$

$\Longrightarrow$ $(X - \mu_x)(Y - \mu_y)$ の期待値, $g(X, Y) = (X - \mu_x)(Y - \mu_y)$

$$\mathrm{Cov}(X, Y)$$
$$= \mathrm{E}((X - \mu_x)(Y - \mu_y))$$

$$
= \begin{cases} \displaystyle\sum_i \sum_j (x_i - \mu_x)(y_j - \mu_y) f(x_i, y_j), \\ \hfill \text{離散型確率変数} \\ \displaystyle\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mu_x)(y - \mu_y) f(x, y) \mathrm{d}x \mathrm{d}y, \\ \hfill \text{連続型確率変数} \end{cases}
$$

いくつかの公式：

1. 確率変数 X, Y について，

   定理：　$\mathrm{E}(X + Y) = \mathrm{E}(X) + \mathrm{E}(Y)$

   証明：

$$
\mathrm{E}(X + Y) = \sum_i \sum_j (x_i + y_j) f(x_i, y_j)
$$

$$
\begin{aligned}
&= \sum_i \sum_j x_i f(x_i, y_j) \\
&+ \sum_i \sum_j y_j f(x_i, y_j) \\
&= \mathrm{E}(X) + \mathrm{E}(Y)
\end{aligned}
$$

2. 確率変数 X と Y が独立のとき，

定理： $\mathrm{E}(XY) = \mathrm{E}(X)\mathrm{E}(Y)$

証明：

$$
\begin{aligned}
\mathrm{E}(XY) &= \sum_i \sum_j x_i y_j f(x_i, y_j) \\
&= \sum_i \sum_j x_i y_j f(x_i) h(y_j)
\end{aligned}
$$

$$= \sum_i x_i f(x_i) \sum_j y_j h(y_j)$$
$$= \mathrm{E}(X)\mathrm{E}(Y)$$

3. 確率変数 X, Y について，

定理： $\mathrm{Cov}(X, Y) = \mathrm{E}(XY) - \mathrm{E}(X)\mathrm{E}(Y)$

証明：

$$\begin{aligned}
&\mathrm{Cov}(X, Y) \\
&= \sum_i \sum_j (x_i - \mu_x)(y_j - \mu_y) p_{ij} \\
&= \sum_i \sum_j (x_i y_j - \mu_x y_j - \mu_y x_i + \mu_x \mu_y) p_{ij} \\
&= \sum_i \sum_j x_i y_j f(x_i, y_j)
\end{aligned}$$

$$
\begin{aligned}
&- \boldsymbol{\mu_x} \sum_i \sum_j \boldsymbol{y_j f(x_i, y_j)} \\
&- \boldsymbol{\mu_y} \sum_i \sum_j \boldsymbol{x_i f(x_i, y_j)} \\
&+ \boldsymbol{\mu_x \mu_y} \sum_i \sum_j \boldsymbol{f(x_i, y_j)} \\
&= \sum_i \sum_j \boldsymbol{x_i y_j f(x_i, y_j) - \mu_x \mu_y - \mu_y \mu_x + \mu_x \mu_y} \\
&= \sum_i \sum_j \boldsymbol{x_i y_j f(x_i, y_j) - \mu_x \mu_y} \\
&= \mathbf{E}(\boldsymbol{XY}) - \mathbf{E}(\boldsymbol{X})\mathbf{E}(\boldsymbol{Y})
\end{aligned}
$$

より，一般的な証明：

$$
\mathbf{Cov}(\boldsymbol{X}, \boldsymbol{Y})
$$

$$
\begin{aligned}
&= \mathrm{E}((\boldsymbol{X} - \boldsymbol{\mu}_x)(\boldsymbol{Y} - \boldsymbol{\mu}_y)) \\
&= \mathrm{E}(\boldsymbol{X}\boldsymbol{Y} - \boldsymbol{\mu}_x\boldsymbol{Y} - \boldsymbol{\mu}_y\boldsymbol{X} + \boldsymbol{\mu}_x\boldsymbol{\mu}_y) \\
&= \mathrm{E}(\boldsymbol{X}\boldsymbol{Y}) - \mathrm{E}(\boldsymbol{\mu}_x\boldsymbol{Y}) - \mathrm{E}(\boldsymbol{\mu}_y\boldsymbol{X}) + \boldsymbol{\mu}_x\boldsymbol{\mu}_y \\
&= \mathrm{E}(\boldsymbol{X}\boldsymbol{Y}) - \boldsymbol{\mu}_x\mathrm{E}(\boldsymbol{Y}) - \boldsymbol{\mu}_y\mathrm{E}(\boldsymbol{X}) + \boldsymbol{\mu}_x\boldsymbol{\mu}_y \\
&= \mathrm{E}(\boldsymbol{X}\boldsymbol{Y}) - \boldsymbol{\mu}_x\boldsymbol{\mu}_y - \boldsymbol{\mu}_y\boldsymbol{\mu}_x + \boldsymbol{\mu}_x\boldsymbol{\mu}_y \\
&= \mathrm{E}(\boldsymbol{X}\boldsymbol{Y}) - \boldsymbol{\mu}_x\boldsymbol{\mu}_y \\
&= \mathrm{E}(\boldsymbol{X}\boldsymbol{Y}) - \mathrm{E}(\boldsymbol{X})\mathrm{E}(\boldsymbol{Y})
\end{aligned}
$$

4. 確率変数 $\boldsymbol{X}$ と $\boldsymbol{Y}$ が独立のとき，

$$\mathrm{E}(\boldsymbol{X}\boldsymbol{Y}) = \mathrm{E}(\boldsymbol{X})\mathrm{E}(\boldsymbol{Y})$$

となるので，

$$\mathrm{Cov}(\boldsymbol{X}, \boldsymbol{Y}) = \boldsymbol{0}$$

を得る。

5. 相関係数 $\boldsymbol{\rho_{xy}}$：

$$\boldsymbol{\rho_{xy}} = \frac{\mathrm{Cov}(\boldsymbol{X}, \boldsymbol{Y})}{\sqrt{\mathrm{V}(\boldsymbol{X})}\sqrt{\mathrm{V}(\boldsymbol{Y})}} = \frac{\mathrm{Cov}(\boldsymbol{X}, \boldsymbol{Y})}{\boldsymbol{\sigma}_x \boldsymbol{\sigma}_y}$$

6. 確率変数 $\boldsymbol{X}$ と $\boldsymbol{Y}$ が独立のとき，

$$\mathrm{Cov}(\boldsymbol{X}, \boldsymbol{Y}) = 0$$

となるので，

$$\boldsymbol{\rho_{xy}} = 0$$

を得る。

7. 確率変数 $\boldsymbol{X}$, $\boldsymbol{Y}$ について，

$$\mathrm{V}(\boldsymbol{X} \pm \boldsymbol{Y}) = \mathrm{V}(\boldsymbol{X}) \pm 2\mathrm{Cov}(\boldsymbol{X}, \boldsymbol{Y}) + \mathrm{V}(\boldsymbol{Y})$$

証明：

$$
\begin{aligned}
&\mathrm{V}(\boldsymbol{X} \pm \boldsymbol{Y}) \\
&= \mathrm{E}(((\boldsymbol{X} \pm \boldsymbol{Y}) - \mathrm{E}(\boldsymbol{X} \pm \boldsymbol{Y}))^2) \\
&= \mathrm{E}(((\boldsymbol{X} - \boldsymbol{\mu}_x) \pm (\boldsymbol{Y} - \boldsymbol{\mu}_y))^2) \\
&= \mathrm{E}((\boldsymbol{X} - \boldsymbol{\mu}_x)^2 \pm 2(\boldsymbol{X} - \boldsymbol{\mu}_x)(\boldsymbol{Y} - \boldsymbol{\mu}_y) \\
&\qquad\qquad + (\boldsymbol{Y} - \boldsymbol{\mu}_y)^2) \\
&= \mathrm{E}((\boldsymbol{X} - \boldsymbol{\mu}_x)^2) \\
&\pm 2\mathrm{E}((\boldsymbol{X} - \boldsymbol{\mu}_x)(\boldsymbol{Y} - \boldsymbol{\mu}_y)) \\
&+ \mathrm{E}((\boldsymbol{Y} - \boldsymbol{\mu}_y)^2) \\
&= \mathrm{V}(\boldsymbol{X}) \pm 2\mathrm{Cov}(\boldsymbol{X}, \boldsymbol{Y}) + \mathrm{V}(\boldsymbol{Y})
\end{aligned}
$$

8. $-1 \leq \boldsymbol{\rho}_{xy} \leq 1$

証明：

次のような t に関する式を考える。$f(t) = \mathrm{V}(\boldsymbol{X}t - \boldsymbol{Y})$

分散なので，必ずゼロ以上となる。よって，すべての t について，$f(t) \geq 0$ となるための条件を求めればよい。t に関する２次方程式の判別式がゼロ以下となる条件を求める。$\mathrm{V}(\boldsymbol{X}t - \boldsymbol{Y}) = \mathrm{V}(\boldsymbol{X}t) - 2\mathrm{Cov}(\boldsymbol{X}t, \boldsymbol{Y}) + \mathrm{V}(\boldsymbol{Y}) = t^2\mathrm{V}(\boldsymbol{X}) - 2t\mathrm{Cov}(\boldsymbol{X}, \boldsymbol{Y}) + \mathrm{V}(\boldsymbol{Y})$

$$\frac{D}{2} = (\mathrm{Cov}(\boldsymbol{X}, \boldsymbol{Y}))^2 - \mathrm{V}(\boldsymbol{X})\mathrm{V}(\boldsymbol{Y}) \leq 0$$

$$\frac{(\mathrm{Cov}(\boldsymbol{X}, \boldsymbol{Y}))^2}{\mathrm{V}(\boldsymbol{X})\mathrm{V}(\boldsymbol{Y})} \leq 1$$

$$-1 \leq \frac{\mathrm{Cov}(\boldsymbol{X}, \boldsymbol{Y})}{\sqrt{\mathrm{V}(\boldsymbol{X})}\sqrt{\mathrm{V}(\boldsymbol{Y})}} \leq 1$$

$$-1 \leq \rho_{xy} \leq 1$$

ρ_{xy} が 1 に近いほど， 正の相関が強くなる。

ρ_{xy} が -1 に近いほど，負の相関が強くなる。

9. 確率変数 X と Y が独立のとき，

   定理： $\mathrm{V}(X+Y) = \mathrm{V}(X) + \mathrm{V}(Y)$

   証明：

   $\mathrm{V}(X+Y) = \mathrm{V}(X) + 2\mathrm{Cov}(X,Y) + \mathrm{V}(Y)$

   確率変数 X と Y が独立のとき，

   $\mathrm{Cov}(X,Y) = 0$

   なので，

   $\mathrm{V}(X+Y) = \mathrm{V}(X) + \mathrm{V}(Y)$

   を得る。

10. n 個の確率変数 X_1, X_2, $\cdots$, X_n について：

    $\mathrm{E}(X_i) = \mu_i$ とするとき，

$\mathbf{E}(\sum_i \boldsymbol{X}_i) = \sum_i \mathbf{E}(\boldsymbol{X}_i) = \sum_i \boldsymbol{\mu}_i$

$\mathbf{V}(\sum_i \boldsymbol{X}_i) \equiv \mathbf{E}(\sum_i (\boldsymbol{X}_i - \boldsymbol{\mu}_i))^2$

$= \mathbf{E}(\sum_i (\boldsymbol{X}_i - \boldsymbol{\mu}_i))(\sum_j (\boldsymbol{X}_j - \boldsymbol{\mu}_j))$

$= \mathbf{E}(\sum_i \sum_j (\boldsymbol{X}_i - \boldsymbol{\mu}_i)(\boldsymbol{X}_j - \boldsymbol{\mu}_j))$

$= \sum_i \sum_j \mathbf{E}((\boldsymbol{X}_i - \boldsymbol{\mu}_i)(\boldsymbol{X}_j - \boldsymbol{\mu}_j))$

$= \sum_i \sum_j \mathbf{Cov}(\boldsymbol{X}_i, \boldsymbol{X}_j)$

11. n 個の確率変数 $X_1, X_2, \cdots, X_n$ は互いに独立で同じ平均 μ，分散 σ^2 を持つとする。すなわち，すべての $i = 1, 2, \cdots, n$ について，

$$\mathbf{E}(\boldsymbol{X}_i) = \boldsymbol{\mu}, \quad \mathbf{V}(\boldsymbol{X}_i) = \boldsymbol{\sigma}^2$$

を仮定する。

さらに，算術平均 $\overline{X} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} X_i$ を考える。

このとき，

定理： $\mathrm{E}(\overline{\boldsymbol{X}}) = \boldsymbol{\mu}$, $\mathrm{V}(\overline{\boldsymbol{X}}) = \dfrac{\boldsymbol{\sigma}^2}{\boldsymbol{n}}$

が成り立つ。

証明：

$$\begin{aligned}
\mathrm{E}(\overline{\boldsymbol{X}}) &= \mathrm{E}(\sum_i \frac{\boldsymbol{X}_i}{\boldsymbol{n}}) = \sum_i \mathrm{E}(\frac{\boldsymbol{X}_i}{\boldsymbol{n}}) \\
&= \sum_i \frac{1}{\boldsymbol{n}} \mathrm{E}(\boldsymbol{X}_i) = \sum_i \frac{1}{\boldsymbol{n}} \boldsymbol{\mu} \\
&= \boldsymbol{\mu}
\end{aligned}$$

$$\begin{aligned}
\mathrm{V}(\overline{\boldsymbol{X}}) &= \mathrm{V}(\sum_i \frac{\boldsymbol{X}_i}{\boldsymbol{n}}) = \sum_i \mathrm{V}(\frac{\boldsymbol{X}_i}{\boldsymbol{n}}) \\
&= \sum_i \frac{1}{\boldsymbol{n}^2} \mathrm{V}(\boldsymbol{X}_i) = \sum_i \frac{1}{\boldsymbol{n}^2} \boldsymbol{\sigma}^2
\end{aligned}$$

$$= \frac{\sigma^2}{n}$$

3.2 いくつかの分布の平均・分散

ベルヌイ分布の平均と分散： **ベルヌイ分布：**

$$f(x) = p^x(1-p)^{1-x} \qquad x = 0, 1$$

$$\mathrm{E}(X) = p,\ \mathrm{V}(X) = p(1-p)$$

証明：

平均：

$$\begin{aligned} \mathrm{E}(X) &= \sum_x x f(x) \\ &= \sum_{x=0}^{1} x p^x (1-p)^{1-x} \end{aligned}$$