2 項分布（Binomial Distribution）の平均と分散：　2 項分布：

$$
\begin{aligned}
f(x) &= {}_nC_x p^x (1-p)^{n-x} \\
&= \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}, \qquad x = 0, 1, 2, \cdots, n,
\end{aligned}
$$

確率関数の性質より，

$$
\sum_{x=0}^{n} {}_nC_x p^x (1-p)^{n-x} = 1
$$

を得る（今回は，和が 1 になることを前提とする）。

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

証明：

平均：

$$
\begin{aligned}
\mu = \mathbf{E}(X) &= \sum_{x=0}^{n} x f(x) = \sum_{x=1}^{n} x f(x) \\
&= \sum_{x=1}^{n} x\, {}_nC_x p^x (1-p)^{n-x} \\
&= \sum_{x=1}^{n} x \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \\
&= \sum_{x=1}^{n} \frac{n!}{(x-1)!(n-x)!} p^x (1-p)^{n-x} \\
&= np \sum_{x=1}^{n} \frac{(n-1)!}{(x-1)!((n-1)-(x-1))!} p^{x-1} (1-p)^{(n-1)-(x-1)} \\
&= np \sum_{x'=0}^{n'} \frac{n'!}{x'!(n'-x')!} p^{x'} (1-p)^{n'-x'}
\end{aligned}
$$

$$= np \sum_{x'=0}^{n'} {}_{n'}C_{x'} p^{x'} (1-p)^{n'-x'}$$
$$= np$$

ただし，$n' = n - 1$，$x' = x - 1$ と定義する。

$x = 1, \mathbf{2}, \cdots, n$ のとき，$x' = 0, \mathbf{1}, \cdots, n'$ に注意。

分散：

$\mathbf{V}(X) = \mathbf{E}(X^2) - \mu^2$ により，$\mathbf{E}(X^2)$ を求める。

$X^2 = X(X-1) + X$ を利用する。

$$\mathbf{E}(X^2) = \mathbf{E}(X(X-1)) + \mathbf{E}(X)$$

したがって，

$\mathbf{V}(X) = \mathbf{E}(X(X-1)) + \mu - \mu^2$ となる。

右辺第 1 項を求める。

$$
\begin{aligned}
\mathbf{E}(X(X-1)) &= \sum_{x=0}^{n} x(x-1)f(x) = \sum_{x=2}^{n} x(x-1)f(x) \\
&= \sum_{x=2}^{n} x(x-1)_n C_x p^x (1-p)^{n-x} \\
&= \sum_{x=2}^{n} x(x-1)\frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \\
&= \sum_{x=2}^{n} \frac{n!}{(x-2)!(n-x)!} p^x (1-p)^{n-x} \\
&= n(n-1)p^2 \sum_{x=2}^{n} \frac{(n-2)!}{(x-2)!(n-x)!} p^{x-2} (1-p)^{n-x}
\end{aligned}
$$

$$
\begin{aligned}
&= n(n-1)p^2 \sum_{x=2}^{n} \frac{(n-2)!}{(x-2)!((n-2)-(x-2))!} p^{x-2}(1-p)^{(n-2)-(x-2)} \\
&= n(n-1)p^2 \sum_{x'=0}^{n'} \frac{n'!}{x'!(n'-x')!} p^{x'}(1-p)^{n'-x'} \\
&= n(n-1)p^2 \sum_{x'=0}^{n'} {}_{n'}C_{x'}\, p^{x'}(1-p)^{n'-x'} \\
&= n(n-1)p^2
\end{aligned}
$$

途中で，$n' = n - 2$**，**$x' = x - 2$ **と定義する。**

まとめると，

$$
\begin{aligned}
\mathbf{V}(X) &= \mathbf{E}(X^2) - \mu^2 \\
&= \mathbf{E}(X(X-1)) + \mu - \mu^2
\end{aligned}
$$

$$= n(n-1)p^2 + np - n^2p^2$$

$$= -np^2 + np$$

$$= np(1-p)$$

ポアソン分布（Poisson Distribution）の平均と分散：　ポアソン分布：

$$f(x) = \frac{e^{-\lambda}\lambda^x}{x!} \qquad x = 0, 1, 2, \cdots$$

ただし，$\lambda > 0$ **とする。確率関数の性質より，**

$$\sum_{x=0}^{\infty} \frac{e^{-\lambda}\lambda^x}{x!} = 1$$

を得る（今回は，和が 1 になることを前提とする）。

$\mathbf{E}(X) = \lambda$，$\mathbf{V}(X) = \lambda$

証明：

平均：

$$\begin{aligned}
\mu = \mathbf{E}(X) = \sum_{x=0}^{\infty} xf(x) &= \sum_{x=1}^{\infty} xf(x) \\
&= \sum_{x=1}^{\infty} x\frac{e^{-\lambda}\lambda^x}{x!} \\
&= \sum_{x=1}^{\infty} \frac{e^{-\lambda}\lambda^x}{(x-1)!} \\
&= \lambda \sum_{x=1}^{\infty} \frac{e^{-\lambda}\lambda^{x-1}}{(x-1)!}
\end{aligned}$$

$$
\begin{aligned}
&= \lambda \sum_{x'=0}^{\infty} \frac{e^{-\lambda} \lambda^{x'}}{x'!} \\
&= \lambda
\end{aligned}
$$

分散：2 項分布の分散の計算と同様に，$\mathbf{V}(X) = \mathbf{E}(X^2) - \mu^2$ により，$\mathbf{E}(X^2)$ を求める。

$X^2 = X(X-1) + X$ **を利用する。**

$\mathbf{E}(X^2) = \mathbf{E}(X(X-1)) + \mathbf{E}(X)$

したがって，

$\mathbf{V}(X) = \mathbf{E}(X(X-1)) + \mu - \mu^2$ **となる。**

右辺第 1 項を求める。

$$
\begin{aligned}
\mathbf{E}(X(X-1)) &= \sum_{x=0}^{n} x(x-1)f(x) = \sum_{x=2}^{n} x(x-1)f(x) \\
&= \sum_{x=2}^{\infty} x(x-1)\frac{e^{-\lambda}\lambda^{x}}{x!} \\
&= \sum_{x=2}^{\infty} \frac{e^{-\lambda}\lambda^{x}}{(x-2)!} \\
&= \lambda^2 \sum_{x=2}^{\infty} \frac{e^{-\lambda}\lambda^{x-2}}{(x-2)!} \\
&= \lambda^2 \sum_{x'=0}^{\infty} \frac{e^{-\lambda}\lambda^{x'}}{x'!} \\
&= \lambda^2
\end{aligned}
$$

よって，$\mathbf{V}(X) = \mathbf{E}(X(X-1)) + \mu - \mu^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$ **となる。**

$\mu = \mathbf{E}(X) = \lambda$ に注意。

4.3　積率 (P.56)

確率変数 X，定数 a

$$\mathbf{E}((X-a)^k)$$

を a の回りの k 次の積率（または，モーメント）と呼ぶ。

通常，$a = \mu = \mathbf{E}(X)$ が使われ，

$$\mathbf{E}((X-\mu)^k)$$

を平均の回りの k 次の積率（または，モーメント）と呼ぶ。

$Z = \dfrac{X - \mu}{\sigma}$ として標準化（または，基準化）する。

X が正規分布であれば，Z の 3 次の積率は $\mathrm{E}(Z^3) = 0$，4 次の積率は $\mathrm{E}(Z^4) = 3$ となる（証明略）。

$\mathrm{E}(Z^3)$ を歪み（歪度），$\mathrm{E}(Z^4)$ を尖り（尖度）と呼ぶ。

$\mathrm{E}(Z^3) < 0$ のとき，X の分布は左に歪んでいる（左のすそ野が広い）となる（正規分布は左右対称で，歪度はゼロ）。

$\mathrm{E}(Z^4) > 3$ のとき，X の分布は正規分布より尖っている（分布の中央が尖っているということは，すそ野が厚いということ）となる。

実際に観測されたデータが正規分布に近いかどうかをみる指標。

この尺度は，昔からある概念ではあるが，データ分析で最近よく使われるようになっている。

4.4　同時確率分布 (P.57)

２つのサイコロ投げで，出る目の数を X, Y とする。

X と Y は独立なので，$X = i, Y = j\ (i, j = 1, 2, 3, 4, 5, 6)$ となる確率は，

$$P(X = i, Y = j) = P(X = i)P(Y = j) = \frac{1}{36}$$

となる。

$\Longrightarrow$ **この確率の系列を確率変数 X と Y の同時確率分布と呼ぶ**

注) 事象 A, B は独立 (P.38):

$$P(A \cap B) = P(A)P(B)$$

一般的に，離散型確率変数 X, Y の同時確率分布は，

$$P(X = x_i, Y = y_j) = p_{ij} = f(x_i, y_j)$$

$i = 1, 2, \cdots, n$, $j = 1, 2, \cdots, m$

によって与えられる。

$f(x_i, y_j)$： **同時確率関数**

$X \setminus Y$	y_1	y_2	$\cdots$	y_m
x_1	p_{11}	p_{12}	$\cdots$	p_{1m}
x_2	p_{21}	p_{22}	$\cdots$	p_{2m}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_n	p_{n1}	p_{n2}	$\cdots$	p_{nm}

たて，よこでそれぞれ総和する。