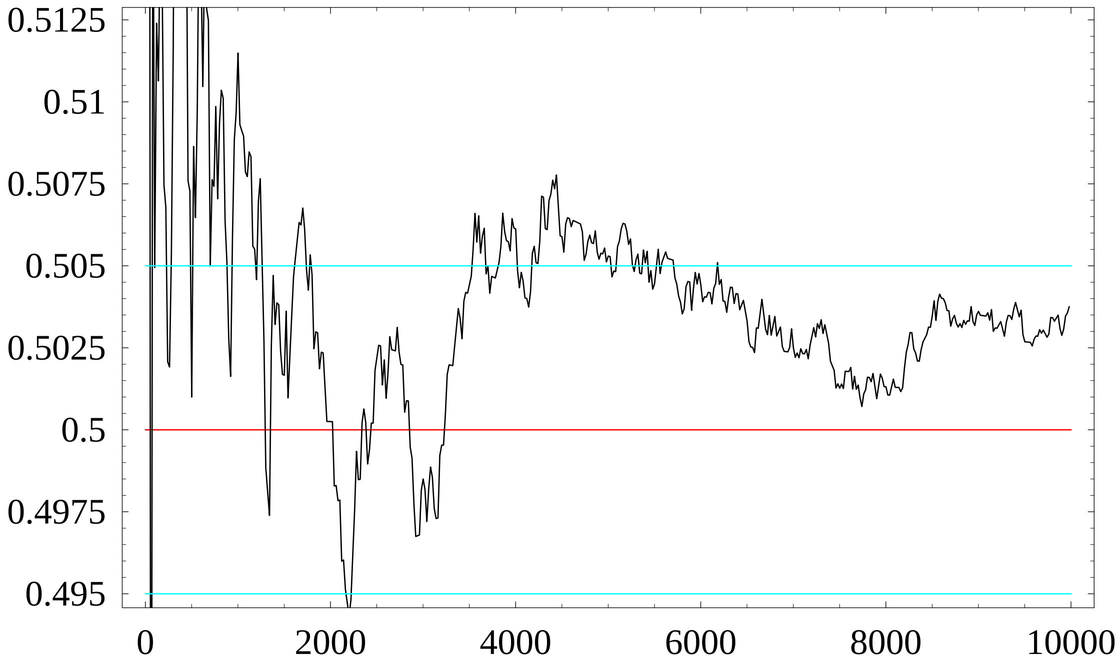
# 第12講：大数の法則・中心極限定理

**大数の弱法則: コインを投げる場合**

正しいコインを、**独立に** に$n = 1, 2, \cdots, 10000$　回投げた場合




横軸：投げる回数；縦軸：表の出る相対頻度

**大数の弱法則**

**大数の弱法則** weak law of large numbers

1. 独立性：確率変数 $X_1, X_2, \cdots, X_n$ が互いに独立
2. 平均の同一性：$\mu = E(X_i)\,, i = 1, 2, \cdots, n$
3. 分散の有限性：$\sigma_i^2 = V(X_i) \leq \sigma^2\,, i = 1, 2, \cdots, n$

このとき、任意の $\epsilon > 0$ に対して

$$\lim_{n\to\infty} P\left(\left|\frac{X_1 + X_2 + \cdots + X_n}{n} - \mu\right| \geq \epsilon\right) = 0$$

このとき、$\bar{X}$ が $\mu$ に **確率収束** converge in probability という。

## 大数の弱法則: 証明

$$Y = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

とおくと、$E(Y) = \mu$ となる。また独立性より

$$V(Y) = \frac{\sigma_1^2 + \cdots + \sigma_n^2}{n^2} \leq \frac{\sigma^2 + \cdots + \sigma^2}{n^2} = \frac{\sigma^2}{n}$$

を得る。チェビシェフの不等式

$$P\left(|Y - E(Y)| \geq \epsilon\right) \leq \frac{V(Y)}{\epsilon^2}$$

より

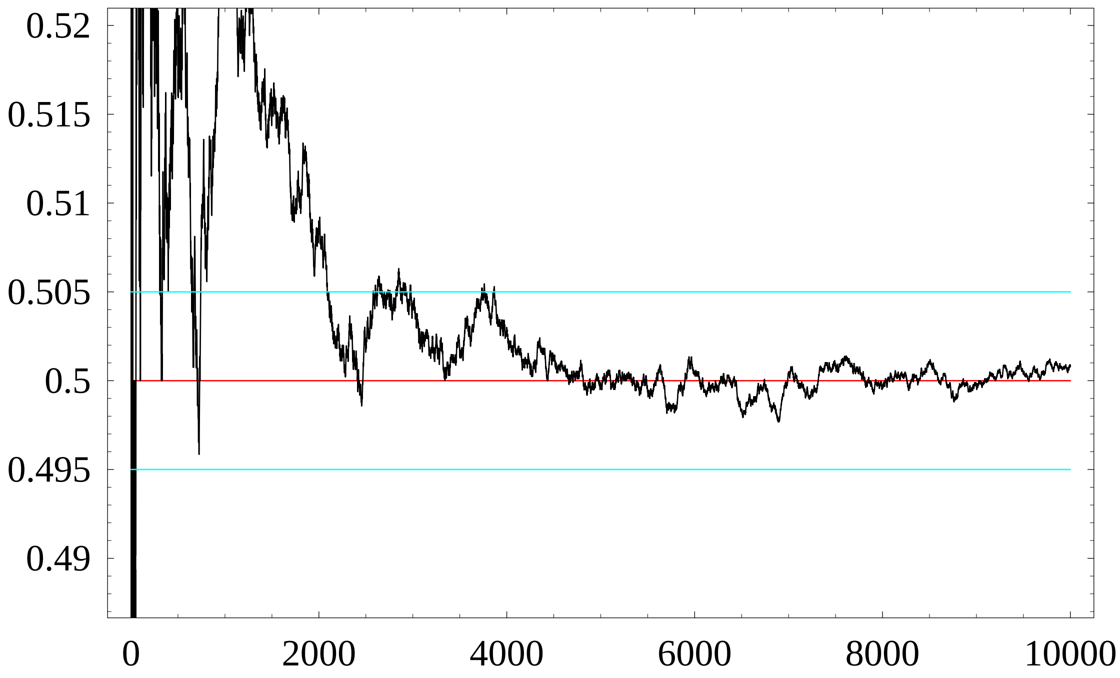
$$P\left(|Y - \mu| \geq \epsilon\right) \leq \frac{\sigma^2}{n\epsilon^2}$$

を得る。したがって、

$$\lim_{n\to\infty} P\left(\left|\frac{X_1 + X_2 + \cdots + X_n}{n} - \mu\right| \geq \epsilon\right) = 0$$

### 大数の強法則: コインを投げる場合

**次々に** 正しいコインを $n = 10000$ 回投げた場合
横軸：投げる回数；縦軸：表の出る相対頻度



## チェビシェフの不等式の拡張

$X$ 平均を $\mu = E(X)$, 偶数次の中心モーメント

$$\nu_{2k} = E(X-\mu)^{2k}$$

とする.

任意の $\epsilon > 0$ に対して、次が成り立つ

$$P\left(|X-\mu| \geq \epsilon\right) \leq \frac{\nu_{2k}}{\epsilon^{2k}}$$

---

**証明:** $D = \{x | x-\mu| \geq \epsilon\}$ とする。

$$\begin{aligned}
\nu_{2k} &= \int_{-\infty}^{\infty} (x-\mu)^{2k} f(x)\,dx \\
&\geq \int_D (x-\mu)^{2k} f(x)\,dx \\
&\geq \int_D \epsilon^{2k} f(x)\,dx \\
&\geq \epsilon^{2k} \int_D f(x)\,dx \\
&\geq \epsilon^{2k} P\left(|X-\mu| \geq \epsilon\right)
\end{aligned}$$

## 標本平均の 4 次のモーメント

- $X_1, X_2, \cdots, X_n$: 互いに独立
- $E(X_i) = \mu$, $V(X_i) = \sigma^2$, $E(X_i - \mu)^4 = \nu_4$, $i = 1, \cdots, n$

$$E(\bar{X} - \mu)^4 = \frac{1}{n^2}\left[\frac{1}{n}\nu_4 + 3(1 - \frac{1}{n})\,\sigma^4\right]$$

---

**証明:**

$$\begin{aligned}
(\bar{X} - \mu)^4 &= \left[\frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)\right]^4 \\
&= \frac{1}{n^4}\left\{\sum_{i=1}^{n}(X_i - \mu)^4 + \sum_{i \neq j} {}_4C_2\,(X_i - \mu)^2(X_j - \mu)^2\right. \\
&\quad + \sum_{i \neq j \neq k \neq \ell} {}_4C_1\,(X_i - \mu)(X_j - \mu)(X_k - \mu)(X_\ell - \mu) \\
&\quad \left. + \sum_{i \neq j} {}_4C_1\,(X_i - \mu)^3(X_j - \mu)\right\}
\end{aligned}$$

両辺期待値を取ると, $E(\bar{X} - \mu)^4 = \frac{1}{n^4}(n\nu_4 + {}_nC_2\ {}_4C_2\,\sigma^4)$

## 大数の強法則

**大数の強法則** strong law of large numbers

1. 確率変数 $X_1, X_2, \cdots, X_n$：互いに独立で同一分布に従う
2. さらに、

$$\begin{cases} \mu = E(X_i) \\ \sigma^2 = V(X_i) \qquad\qquad i = 1, 2, \cdots, n \\ \nu_4 = E(X_i - \mu)^4 \end{cases}$$

このとき、

$$P\left(\lim_{n\to\infty} \frac{X_1 + X_2 + \cdots + X_n}{n} = \mu\right) = 1$$

このとき、$\bar{X}$ が $\mu$ に **概収束** converge almost surely (converge with probability 1) という。

## 大数の強法則の証明

- 評価すべき事象は
$$\lim_{n\to\infty} \bar{X} = \mu$$
である。

- この事象は次の事象と同等である。任意の $\epsilon > 0$ に対して、自然数 $N_\epsilon$ が存在し, 全ての $n > N_\epsilon$ に対して
$$\left|\bar{X} - \mu\right| < \epsilon\,, \quad \forall\epsilon > 0,\, n > N_\epsilon$$
という事象である。

- この事象の余事象は、**ある** $\epsilon$ **が存在し、この** $\epsilon$ **に対して、どんな大きい** $N$ **をとっても**, $n > N$ **が存在し**、次が満たされる事象
$$\left|\bar{X} - \mu\right| \geq \epsilon\,, \quad \text{for some } n > N$$
である。

## 大数の強法則の証明（つづき）

一方、$n > N > \nu_4$ のとき、チェビシェフの不等式より

$$\begin{aligned} P_n &= P\left(|\bar{X}-\mu| \ge \epsilon\right) \\ &\le \frac{E(\bar{X}-\mu)^4}{\epsilon^4} \\ &= \frac{1}{n^2\epsilon^4}\left[\frac{1}{n}\nu_4 + 3(1-\frac{1}{n})\,\sigma^4\right] \\ &< \frac{1+3\sigma^4}{\epsilon^4}\,\frac{1}{n^2} \end{aligned}$$

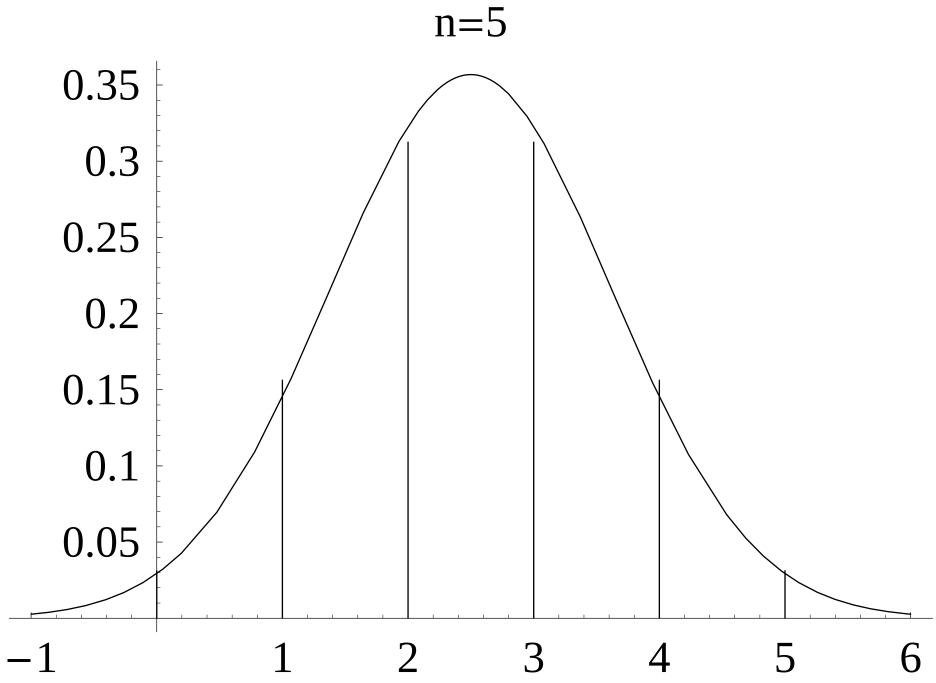
したがって, **1つの** $\epsilon > 0$ と**任意**の $N$ に対して、

$$\begin{aligned} P\left(\lim_{n\to\infty}\bar{X}=\mu\right) &= 1 - P\left(|\bar{X}-\mu| \ge \epsilon| \text{ for some } n > N\right) \\ &\ge 1 - (P_N + P_{N+1} + \cdots) \\ &\ge 1 - \frac{1+3\sigma^4}{\epsilon^4}\left\{\frac{1}{N^2} + \frac{1}{(N+1)^2} + \cdots +\right\} \\ &\stackrel{N\to\infty}{\longrightarrow} 1 \end{aligned}$$

最後の極限は、無限級数 $\sum_{i=1}^{\infty} 1/N^2$ の収束性による。

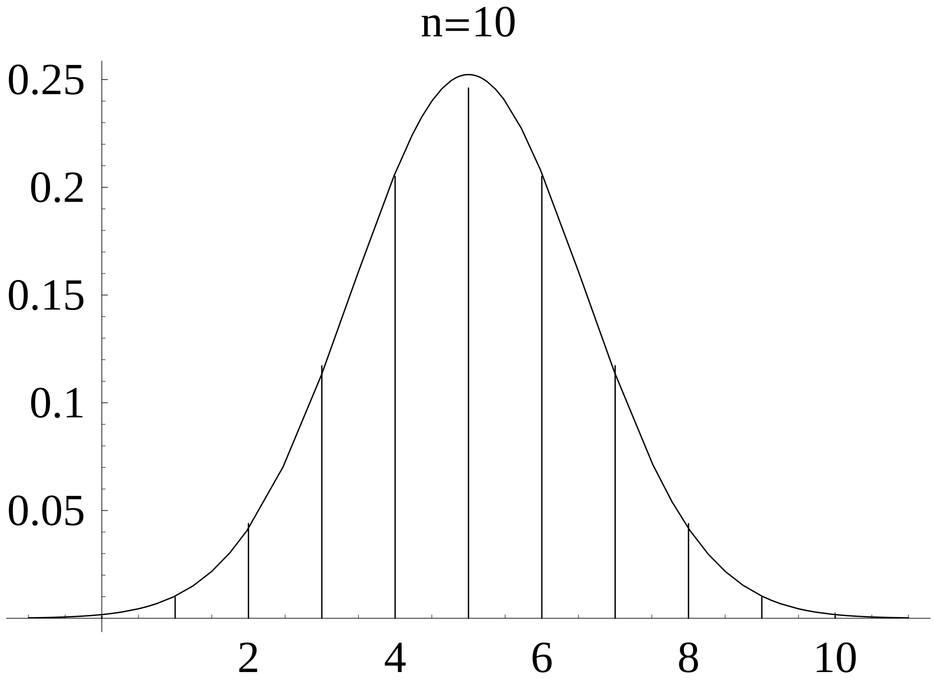
## 中心極限定理: 二項分布の場合

$X_1, \cdots, X_n$: 独立で、$p = 1/2$ の Bernoulli 分布に従い、$\sum_{i=1}^{n} X_i$ の分布



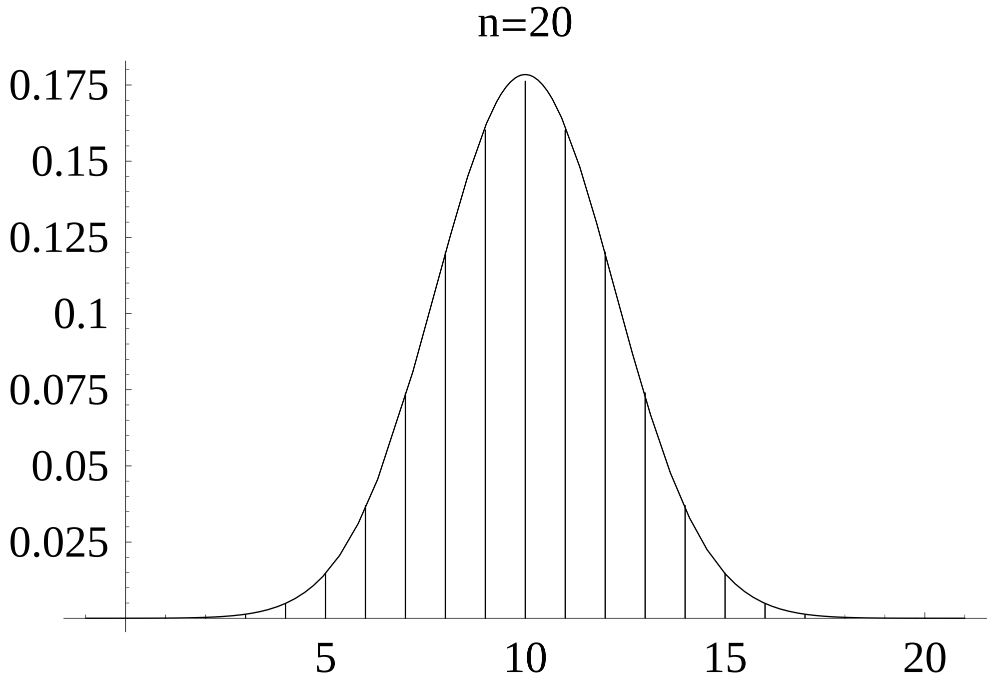
**中心極限定理: 二項分布の場合（つづき）**

$X_1, \cdots, X_n$: 独立で、$p = 1/2$ の Bernoulli 分布に従い、$\sum_{i=1}^{n} X_i$ の分布



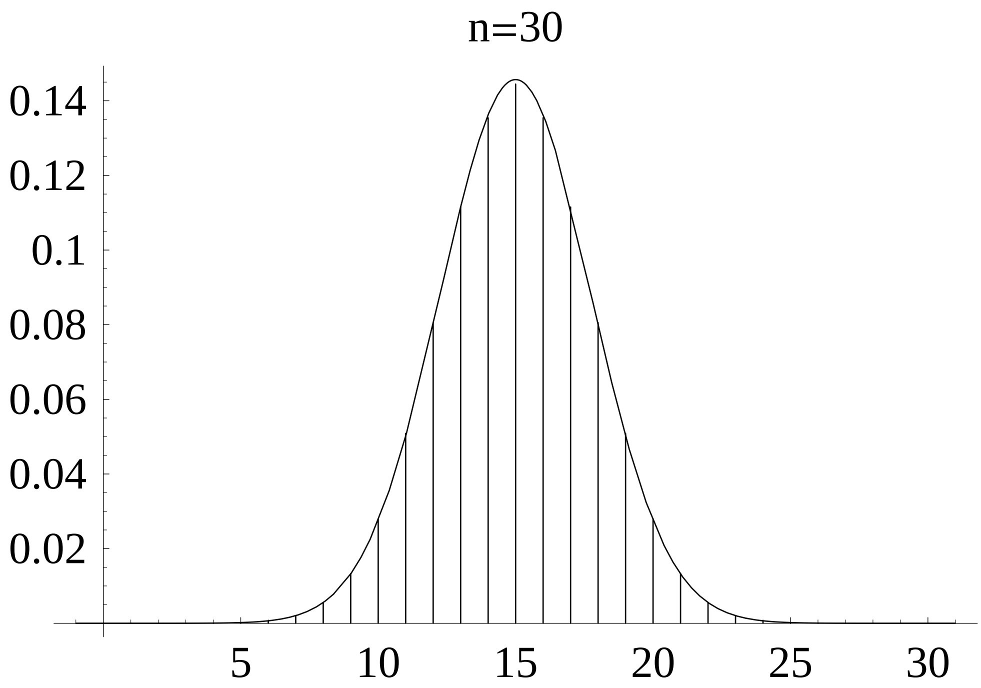
## 中心極限定理: 二項分布の場合（つづき）

$X_1, \cdots, X_n$: 独立で、$p = 1/2$ の Bernoulli 分布に従い、$\sum_{i=1}^{n} X_i$ の分布



## 中心極限定理: 二項分布の場合（つづき）

$X_1, \cdots, X_n$: 独立で、$p = 1/2$ の Bernoulli 分布に従い、$\sum_{i=1}^n X_i$ の分布



## 中心極限定理 Central Limit Theorem

**定理 1** 次の条件の下で

1. $X_1, \cdots, X_n$ が独立で、同じ分布に従う
2. $E(X_i) = \mu,\ V(X_i) = \sigma^2,\ i = 1, \cdots, n$

確率変数 $Y = \frac{\sqrt{n}(\bar{X}-\mu)}{\sigma}$ は標準正規分布に弱収束する。すなわち

$$\lim_{n\to\infty} P\left[\frac{\sqrt{n}(\bar{X}-\mu)}{\sigma} \le y\right] = \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx \tag{1}$$

## 中心極限定理: 証明 (1/2)

**$Y$ の積率母関数が標準正規分布のそれに近づくことを証明する。**

$X_i$ の積率母関数が存在するという、定理より強い条件を仮定する。

$Y_i = \frac{X_i - \mu}{\sigma}$ とすると、

$$E(Y_i) = 0,\ V(Y_i) = 1,\ Y = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} Y_i$$

$Y_i$ の積率母関数を $M(t)$ とすると

$$\begin{aligned} M(0) &= E\left(e^{0\,Y_i}\right) = 1 \\ M'(0) &= E(Y_i) = 0 \\ M''(0) &= E(Y_i^2) = V(Y_i) + (EY_i)^2 = 1 \end{aligned}$$

また $\psi(t) = \log M(t)$ とすると、

$$\psi'(t) = \frac{M'(t)}{M(t)}, \quad \psi''(t) = \frac{M''(t)M(t) - [M'(t)]^2}{[M(t)]^2}$$

したがって、$\psi(0) = 0,\ \psi'(0) = 0,\ \psi''(0) = 1$ となる。

## 中心極限定理: 証明 (2/2)

一方, ある $|\theta| < |t|$ が存在し、原点の周りで $\psi(t)$ を展開すると

$$\psi(t) = \psi(0) + \frac{\psi'(0)}{1!}t + \frac{\psi''(0)}{2!}t^2 + \frac{\psi'''(\theta)}{3!}t^3 = \frac{1}{2}t^2 + \frac{\psi'''(\theta)}{6}t^3$$

従って、$Y$ の積率母関数は

$$\begin{aligned}
E\left(e^{Yt}\right) &= E\left(e^{\frac{1}{\sqrt{n}}\sum_{i=1}^{n} Y_i t}\right) \\
&= \prod_{i=1}^{n} E\left(e^{\frac{t}{\sqrt{n}}Y_i}\right) \\
&= \left[M\left(t/\sqrt{n}\right)\right]^n \\
&= \exp\left\{n \log M\left(t/\sqrt{n}\right)\right\} \\
&= \exp\left\{n\left[\frac{1}{2}\left(t/\sqrt{n}\right)^2 + \frac{\psi'''(\theta)}{6}\left(t/\sqrt{n}\right)^3\right]\right\} \\
&= \exp\left\{\frac{1}{2}t^2 + \frac{\psi'''(\theta)}{6}\frac{t^3}{\sqrt{n}}\right\} \\
&\xrightarrow{n\to\infty} \exp\left\{\frac{1}{2}t^2\right\}
\end{aligned}$$

**二項分布の正規近似**

**定理 2 (De Moivre-Laplace Limit Theorem)**
$X$ が二項分布に従うならば

$$X \sim \mathrm{Bi}(n,\, p)$$

次が成り立つ

$$\lim_{n\to\infty} P\left[a \leq \frac{X - np}{\sqrt{np(1-p)}} \leq b\right] = \Phi(b) - \Phi(a) \tag{2}$$

ただし、$\Phi(x)$ は標準正規分布の分布関数である。すなわち、

$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2}\, dy$$

## 二項分布の正規近似: 証明

$X_1, \cdots, X_n$ を独立で、成功する確率が $p$ の Bernoulli 分布に従うならば,

$$\begin{aligned} X &= \sum_{i=1}^{n} X_i \sim \mathrm{Bi}(n,\, p) \\ E(X_i) &= p \\ V(X_i) &= p(1-p) \end{aligned}$$

となる。
中心極限定理によって

$$\begin{aligned} \frac{X - np}{\sqrt{np(1-p)}} &= \frac{\sum_{i=1}^{n} X_i - np}{\sqrt{np(1-p)}} \\ &= \frac{\sqrt{n}(\bar{X} - p)}{\sqrt{p(1-p)}} \\ &\longrightarrow N(0,\, 1) \end{aligned}$$