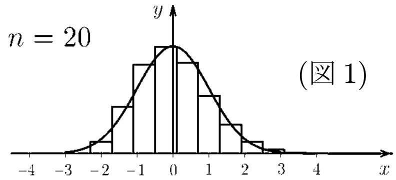
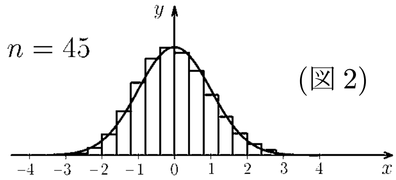
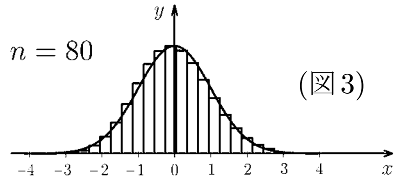
< 中心極限定理 >



(図 1)



(図 2)



(図3)

例　前ページ例 4 の場合 X の平均は $n \times \dfrac{1}{6}$ ，
標準偏差は $\sqrt{n \times \frac{5}{36}}$ より，標準化された確率変数は

$$\boldsymbol{X^*} = \frac{X - (X\text{ の平均})}{(X\text{ の標準偏差})} = \frac{X - n \times \frac{1}{6}}{\sqrt{n \times \frac{5}{36}}}$$

である。
　図 1 は $n = 20$ の場合の $\boldsymbol{X^*}$ の確率分布をヒストグラムの形にしたもので，各長方形の面積の和が 1 になるようにしたものである。このヒストグラムと標準正規分布 $N(0, 1)$ の密度関数 $y = \dfrac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ のグラフを重ねて書いてある。
　図 2 は $n = 45$ の場合であり，図 3 は $n = 80$ の場合である。
　n が大きくなるにつれて $\boldsymbol{X^*}$ の分布は標準正規分布に近づく。すなわち

$$\lim_{n \to \infty} P(a \leqq \boldsymbol{X^*} \leqq b) = \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$$

が任意の $a < b$ に対して成り立つ。
　一方 22 ページ例 1 のように X_k を定めると $X_1, X_2, \cdots, X_n$ は独立で，しかも同じ分布 $\left(E[X_k] = \dfrac{1}{6}, V\big(X_k\big) = \dfrac{5}{36} \right)$ であり，

$$X = X_1 + X_2 + \cdots + X_n$$

と表される。従って上の極限式は次のようになる。$\left(\text{ただし} (\boldsymbol{X_1 + X_2 + \cdots + X_n})^* = \dfrac{X_1 + X_2 + \cdots + X_n - n \times \frac{1}{6}}{\sqrt{n \times \frac{5}{36}}} \right)$

$$\lim_{n \to \infty} P\Big(a \leqq (\boldsymbol{X_1 + X_2 + \cdots + X_n})^* \leqq b \Big) = \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$$

　一般に確率変数列 $X_1, X_2, \cdots, X_n, \cdots$ が独立で同じ確率分布をもつとする。
平均を $m = E[X_k]$，分散を $v = V\big(X_k\big)$ とすると，和 $X_1 + X_2 + \cdots + X_n$ を標準化した確率変数

(1) $$\boxed{(\boldsymbol{X_1 + X_2 + \cdots + X_n})^* = \frac{X_1 + X_2 + \cdots + X_n - nm}{\sqrt{nv}}}$$ (**和の標準化**)

の分布は標準正規分布に近づく。すなわち

(2) $$\boxed{\lim_{n \to \infty} P\Big(a \leqq (\boldsymbol{X_1 + X_2 + \cdots + X_n})^* \leqq b \Big) = \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx}$$ (**中心極限定理**)

が成り立つ (証明略)。これを**中心極限定理**という。(X_k の確率分布がどんな形でも成立する)
ここで (1) 式を (2) に代入し，$a = -M$ ，$b = M$ とおきかえると (2) 式は

(3) $$\lim_{n \to \infty} P\left(\left| \frac{X_1 + X_2 + \cdots + X_n}{n} - m \right| \leqq M \sqrt{\frac{v}{n}} \right) = \int_{-M}^{M} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$$

と書きなおせる。ここで $n \to \infty$ のとき $M\sqrt{\dfrac{v}{n}} \to 0$ に注意してほしい。証明は略すが (3) 式から

(4) $$\lim_{n \to \infty} \frac{X_1 + X_2 + \cdots + X_n}{n} = m$$

が示される。(3) 式は「(4) 式の収束の速さが $\dfrac{1}{\sqrt{n}}$ に比例する」ことを意味している。

問　(1) 式と (2) 式から (3) 式を導け。