統計的な推測【二項分布の正規分布による近似】　p.81〜83

1 学習内容の説明 ⇒ 2 問題演習 ⇒ 3 振り返り(確認テスト・相互採点・リフレクションの記入)

【態度目標】しゃべる、質問する、説明する、動く、協力する、貢献する

【内容目標】二項分布のグラフについて n を大きくして正規分布曲線に近似させる考え方を理解しよう

□二項分布の正規分布による近似

二項分布と正規分布の関連について調べよう。

1 回の試行で事象 A の起こる確率を p とするとき，この試行を n 回行う反復試行において，A の起こる回数を X とすれば，X は二項分布 $B(n,\ p)$ に従う確率変数である。$X=r$ である確率を P_r とすると

$P_r={}_n\mathrm{C}_r p^r q^{n-r}$　　ただし，$q=1-p$，$r=0,\ 1,\ 2,\ \cdots\cdots,\ n$

であり，X の期待値 $E(X)$，標準偏差 $\sigma(X)$ は，次のようになる。

$$E(X)=np,\qquad \sigma(X)=\sqrt{npq}$$

p.72 のおさらい

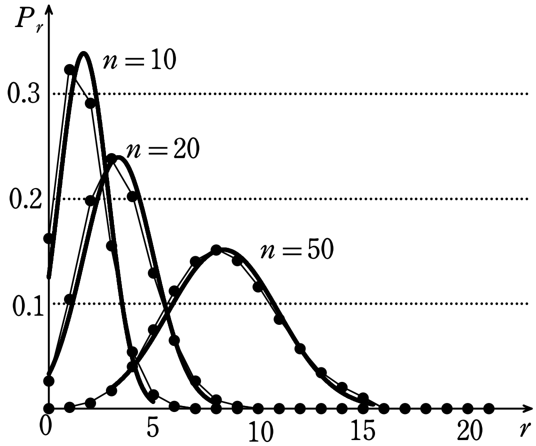
B(回数，確率) なので反復試行の回数が増えていくとどうなるかを考える

二項分布 $B\left(n,\ \dfrac{1}{6}\right)$ に従う確率変数 X について，$n=10,\ 20,\ 50$ のとき，確率 P_r を求めて折れ線グラフをかくと右の図のようになる。

このXと期待値，標準偏差の等しい連続型確率変数が従う正規分布 $N\left(\dfrac{n}{6},\ \dfrac{5n}{36}\right)$ の分布曲線を重ねてかくと右の図のようになる。

この図から，二項分布のグラフの形は，**n が大きくなるにつれて正規分布曲線に近づく**ことが予想される。

また，実際にそうなることが知られている。



二項分布と正規分布　　ゴルトンボード

一般に，次のことが成り立つ。いずれの場合も，$q=1-p$ とする。

二項分布の正規分布による近似

1　二項分布 $B(n,\ p)$ に従う確率変数 X は，n が大きいとき，近似的に正規分布 $N(np,\ npq)$ に従う。

2　二項分布 $B(n,\ p)$ に従う確率変数 X に対し，$Z=\dfrac{X-np}{\sqrt{npq}}$ は，n が大きいとき，近似的に標準正規分布 $N(0,\ 1)$ に従う。

二項分布なので
期待値 $=np$
分散 $=npq$

X の標準化
$Z=\dfrac{X-\text{期待値}}{\text{標準偏差}}$

統計的な推測【二項分布の正規分布による近似】　p.81〜83

例題４） １個のさいころを 720 回投げて，１の目が出る回数を X とするとき，
$110 \leqq X \leqq 130$ となる確率を，標準正規分布 $N(0,\ 1)$ で近似する方法で求めよ。

B(回数，確率)

解答　１の目が出る確率は $\dfrac{1}{6}$ で，X は二項分布 $B\left(720,\ \dfrac{1}{6}\right)$ に従う。

X の期待値 m と標準偏差 σ は

$$m = 720 \cdot \frac{1}{6} = 120, \qquad \sigma = \sqrt{720 \cdot \frac{1}{6} \cdot \frac{5}{6}} = 10$$

二項分布なので
期待値 $= np$
分散 $= npq \Rightarrow$ 標準偏差 $= \sqrt{分散}$

$\therefore$　$B\left(720,\ \dfrac{1}{6}\right)$ に従う確率変数 X は，n が大きいとき，近似的に正規分布 $N(120,\ 10^2)$ に従う。

よって，$Z = \dfrac{X-120}{10}$ は近似的に標準正規分布 $N(0,\ 1)$ に従う。

X の標準化　$Z = \dfrac{X - 期待値}{標準偏差}$

$X = 110$ のとき　$Z = \dfrac{110-120}{10} = -1$

$X = 130$ のとき　$Z = \dfrac{130-120}{10} = 1$

二項分布によるヒストグラムと正規分布に従う確率密度関数を比べると
$P(a \leqq X \leqq b) \fallingdotseq P(a-0.5 \leqq Y \leqq b+0.5)$
と置き換えた方が近似度がよくなる。この方法を半整数補正という。

であるから，求める確率は

$$\begin{aligned} P(110 \leqq X \leqq 130) &= P(-1 \leqq Z \leqq 1) \\ &= 2P(0 \leqq Z \leqq 1) \\ &= 2p(1) = 2 \times 0.3413 = 0.6826 \end{aligned}$$

正規分布表より

研究　連続型確率変数の期待値，分散，標準偏差

連続型確率変数 X のとる値の範囲が $\alpha \leqq X \leqq \beta$ で，その確率密度関数が $f(x)$ であるとき，X の期待値 $m = E(X)$ と分散 $V(X)$ を，次の式で定める。また，X の標準偏差 $\sigma(X)$ は，$\sqrt{V(X)}$ で定める。

連続型確率変数の期待値，分散

期待値　$m = E(X) = \displaystyle\int_\alpha^\beta x f(x)\,dx$

分散　$V(X) = \displaystyle\int_\alpha^\beta (x-m)^2 f(x)\,dx$

離散的な確率変数の期待値，分散は

期待値　$m = E(X) = \displaystyle\sum_{k=1}^{n} x_k p_k$

分散　$V(X) = \displaystyle\sum_{k=1}^{n} (x_k - m)^2 p_k$

離散的な確率変数から連続型の確率変数への
変形には数学Ⅲで学ぶ区分求積法の考えを用いる

たとえば，確率密度関数が

$f(x) = 2x \quad (0 \leqq x \leqq 1)$　７５ページ例１６

である確率変数 X について，期待値 m を計算すると次のようになる。

$$m = E(X) = \int_0^1 (x \cdot 2x)\,dx = \int_0^1 2x^2\,dx = \left[\frac{2}{3}x^3\right]_0^1 = \frac{2}{3}$$

$f(x) = 2x$

また，分散 $V(X)$ は次のように計算される。

$$\begin{aligned} V(X) &= \int_0^1 \left(x - \frac{2}{3}\right)^2 \cdot 2x\,dx \\ &= \int_0^1 \left(x^2 - \frac{4}{3}x + \frac{4}{9}\right) \cdot 2x\,dx = \int_0^1 \left(2x^3 - \frac{8}{3}x^2 + \frac{8}{9}x\right)dx \\ &= \left[\frac{x^4}{2} - \frac{8}{9}x^3 + \frac{4}{9}x^2\right]_0^1 = \frac{1}{18} \end{aligned}$$

$(x-m)^2 f(x) = \left(x - \dfrac{2}{3}\right)^2 \cdot 2x$

よって，標準偏差 $\sigma(X)$ は　$\sigma(X) = \sqrt{\dfrac{1}{18}} = \dfrac{\sqrt{2}}{6}$