動的視覚化法による"目からうろこ"シリーズ

プログラムで理解する視覚でわかる統計確率論（最終回）

二項分布と正規分布に強くなる

浪平 博人

ある確率で起こる出来事について，起きるか，起きないか，という分布は二項分布と呼ばれます．これは基本的な分布の一つで，もの作りの現場でも目にすることがあります．この二項分布が正規分布で近似されることは知られています．今回は，この分布について考えてみましょう．（筆者）

二項分布とは

分布から何が起きるかを実際に観測することを試行といいます．

確率pで起こるある事象Aを取り上げ，1回の試行においてこの事象Aが起きたか起きないかのみに着目しましょう．事象Aの例としては，次のようなものがあります．

- コインを投げて表が出る
- ある病気に感染した患者が発病する
- ある機械が1日壊れないで動く

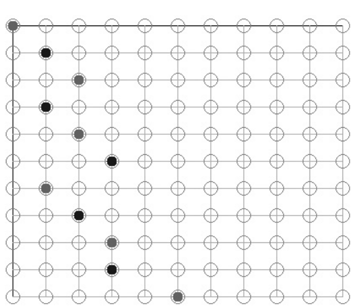
これらの試行をn回繰り返して，事象Aが起きた回数を数えると，0，1，2，…，nのいずれかとなります．事象Aの起きた回数をxとすると，xは0～nまでの範囲で起こり得る確率変数といえます．このxの従う分布を「二項分布」といいます．試行間の結果は互いにほかの試行に影響を与えない（独立である）ことが前提とされています．

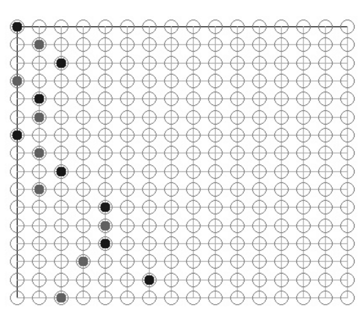
二項分布は，確率分布を考える上での基本的な分布の一つです．この分布の挙動を視覚化してみましょう．

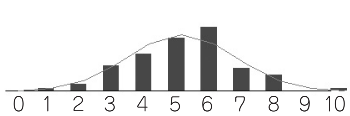
パチンコのモデルで考える

特殊なパチンコを考えます．上から落とされた球は格子状の道をたどりながらn段階を経て下に達するとしましょう．ある段階で，事象Aが起きれば球は確率pで一つ右下の格子に移り，事象Aが起きなければ確率q（$= 1 - p$）で真下へ移動します．最終段階nでの球の右への移動距離は，n回の試行で着目する事象Aの起きた回数を表します．

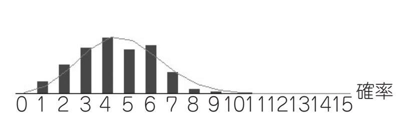
図1（a）は，確率$p = 0.5$，段階$n = 10$と設定して，左上から球を100回落とす試験でのある時点の状態です．一つの球の軌跡の連続的な経過を示すものではありません．それぞれの球が下に落ちるにつれ次第に右に移っていくようすがわかります．**図1（b）**の棒グラフは，100個の球を試行したときの右への移動距離を調べた統計で，同図内の実線は理論分布です．**図1（c）**は，確率$p = 0.3$，$n = 15$として200個の球を落とした場合のシミュレーションです．**図1（d）**は**図1（c）**の実験での移動距離を調べた統計です．**図1（b）**と**図1（d）**を比べると，統計を取る数が多くなれば，実験からの分布は理論分布に安定的に近付いてくることが観測されます．


（a）確率p=0.5,試行回数n=10のモデル


（c）確率p=0.3,試行回数n=15のモデル



（b）（a）での右への移動距離の統計



（d）（c）での右への移動距離の統計

図1 パチンコのモデルで分布を考える

二項分布を表す式

事象Aの起きる確率をpとし，これをn回繰り返したとき事象Aがx回起きる確率$P(x)$は二項分布と呼ばれ，次