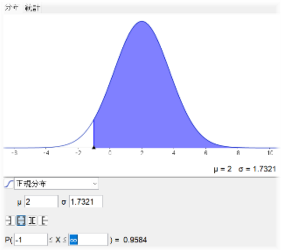
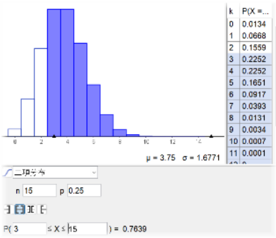
雑感　確率分布と GeoGebra

■　新カリ数Ｂの「統計的な推測」がほぼ必須になって(？)，確率分布・統計的な推測の授業が着々と展開されている（のだろうか）.
　不慣れな内容で，教える側に多くの戸惑いがあるに違いない.

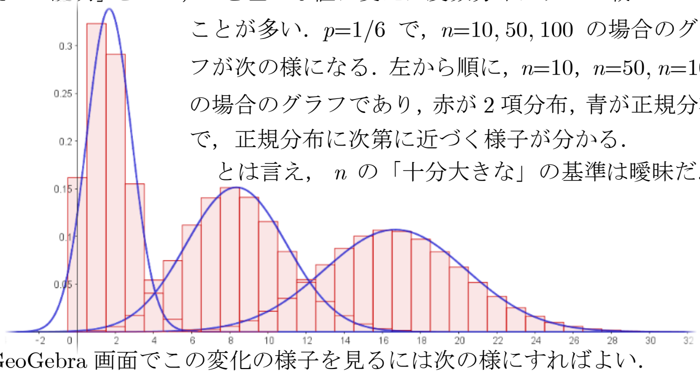
■　この内容の大元にあるのが確率分布である．高校では，2 項分布と正規分布の学習が中心となるが，均一分布をはじめ条件で指定されるような一般の連続型確率分布も定義・基本を押さえておく必要がある（これまでの共通テストでの出題もある；例えば 2022 年には 1 次関数形）.



■　正規分布表の使いかたの説明などで，度数分布グラフが欲しいことがある．厳密な形である必要はないので，フリーハンドで良いのだが，印刷物などではきちんとしたものが望ましいことがある.
　GeoGebra では, 例えば ＞表示＞確率計算器 で簡単に描写できる．例えば正規分布 $N(2,3)$ の $-1 \leqq X$ の範囲であれば，右上のように一発である.
　2 項分布 $B(15, 1/4)$で, $3 \leqq X$ の範囲の図示も右のように同様である.

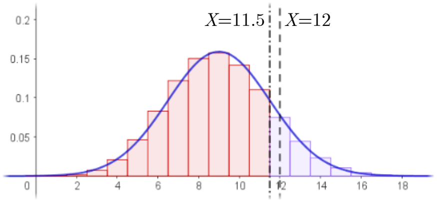


■　授業で必ず取り扱う, 2 項分布の正規分布近似を考えてみる．教科書には,
　2 項分布 $B(n, p)$ に従う確率変数 X は,
　n が十分大きいとき，近似的に正規分布 $N(np, np(1\text{-}p))$ に従う
といったことが，書かれているはずである.
　その「説明」として，n を色々な値に変えた度数分布グラフが載っていることが多い．p=1/6 で，n=10, 50, 100 の場合のグラフが次の様になる．左から順に，n=10，n=50, n=100 の場合のグラフであり，赤が 2 項分布，青が正規分布で，正規分布に次第に近づく様子が分かる.
　とは言え，n の「十分大きな」の基準は曖昧だ.



　GeoGebra 画面でこの変化の様子を見るには次の様にすればよい.
　通常の画面で，n と p のスライダーを作り，範囲を 0<n<150, 0<p<1 程度にして置く．2 項分布は BinomialDist(n, p) で OK である．正規分布は f:Normal(np, sqrt(p(1 - p)n), x, true) とすると，累積分布のグラフが現れるので，さらに f’ とすると，山形の正規分布のグラフが現れる(f は非表示に).
　上の場合は，p=1/6=0.166…と設定し，n のスライダーを動かした.

■　若干レベルの高い話だが, 2 項分布を正規分布近似するとき，半整数補正が精度向上に役立つという話では，右のような図を用意すれば分かり易くなる.



　元の 2 項分布は $B(30,0.3)$ で，これを正規分布近似したものが青い曲線．$P(12 \leqq X)$ は正しくは青い長方形の面積和で，≒0.1593．これを青い曲線の破線 X=12 の右側部分の面積≒0.1160 と近似するのは，値が小さすぎる．これを半整数補正して，1 点破線 X=11.5 の右側の面積≒0.1596 と近似する方が良かろうという話である（2017 年滋賀大に出題がある）.

■　なお，この近似の精度は n だけでなく p にも関係し，p が 0.5 に近ければ，小さな n でも良い近似が得られ，np>5 かつ $n(1\text{-}p)$>5 の場合に適用可能という記述が,『日常のなかの統計学』(鷲尾泰俊；岩波書店) にある.
上の例では np=9>5, $n(1\text{-}p)$=21>5 で，補正すればなかなか良い近似である.

2023 年 09 月 01 日