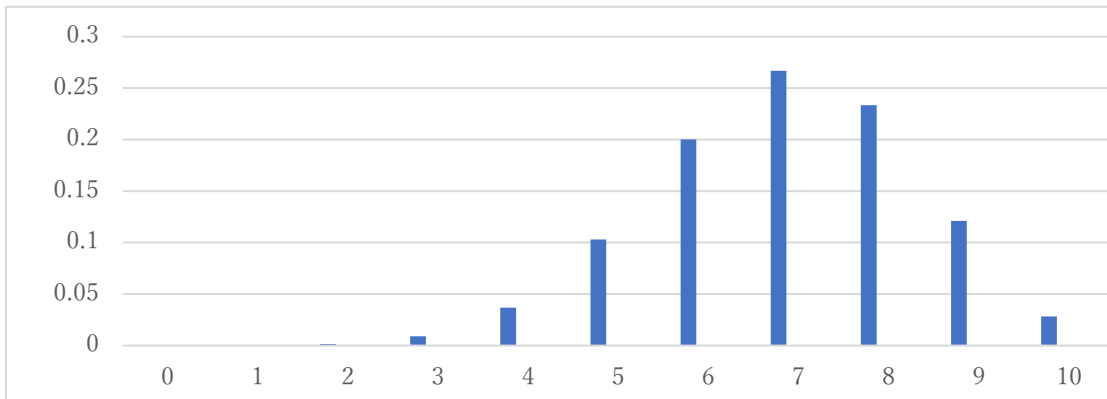


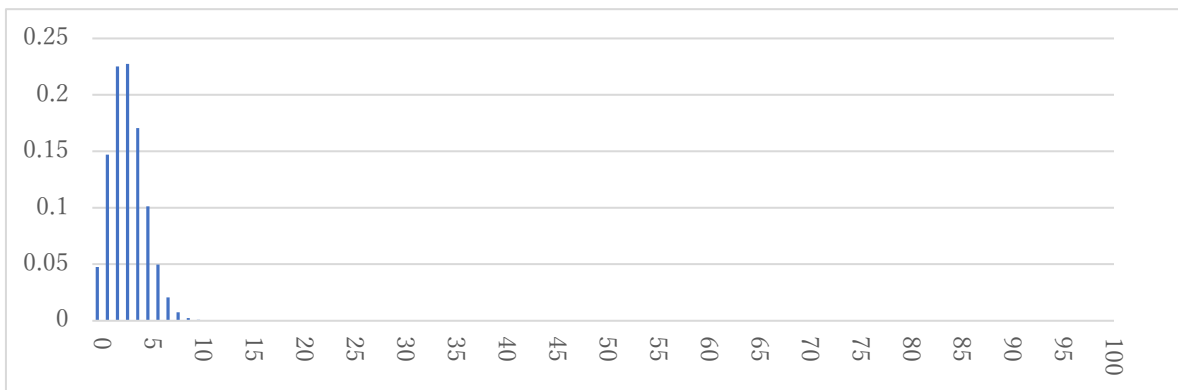
## ・二項分布

高校で「コインを5回投げたときに表が2回出る確率」「当たる確率10%のクジを10回中引いたとき1回当たる確率」などを計算した記憶がありませんか？二項分布は成功する確率が  $p$  の試行を独立に  $n$  回行ったときに成功回数の分布を表現する数学モデルです。簡単な例では、サイコロを10回投げてその中で5の目が出る回数や、コインを20回投げて表の出る回数などがこの分布に従います。例えば、成約率70%の営業マンが10回訪問営業した時の成約数の分布も二項分布で表され、形は次のようになります。90%の確率で成約数は5~9件という計算になります。



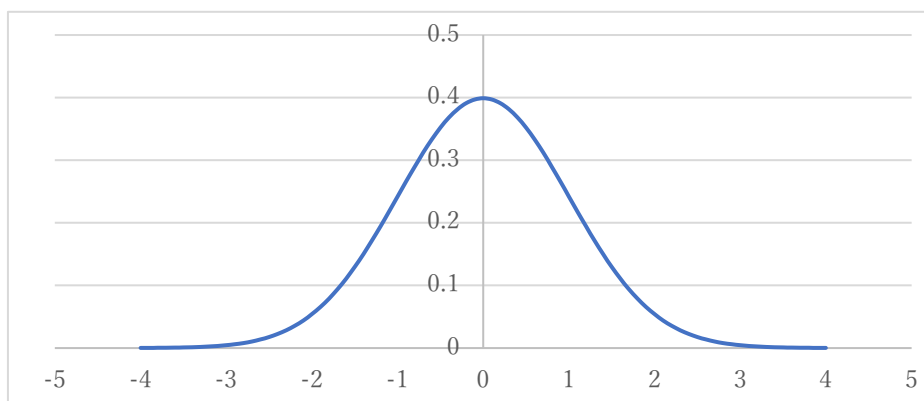
## ・ポアソン分布

ポアソン分布は、めったに起こらない事柄を大量に観察したときに、それが起こる回数の分布を表す数学モデルです。数学的にいうと、二項分布の成功確率  $p$  を限りなく小さく、試行回数  $n$  を限りなく大きくしたときに現れるのがポアソン分布です。馬に蹴られて亡くなった兵士の数の分布はこの分布でよく表されるという有名な話があります。また、あるwebサイトの1時間あたりのアクセス数や、ある面積あたりのチーズフォンデュ専門店数の分布などもこの分布に良くあてはまることが知られています。例えば、成約率3%の営業マンが100回訪問営業した時の成約数の分布は大体ポアソン分布で表され、形は次のようになります。7件以上成約出来る確率はほぼ0という計算になります。



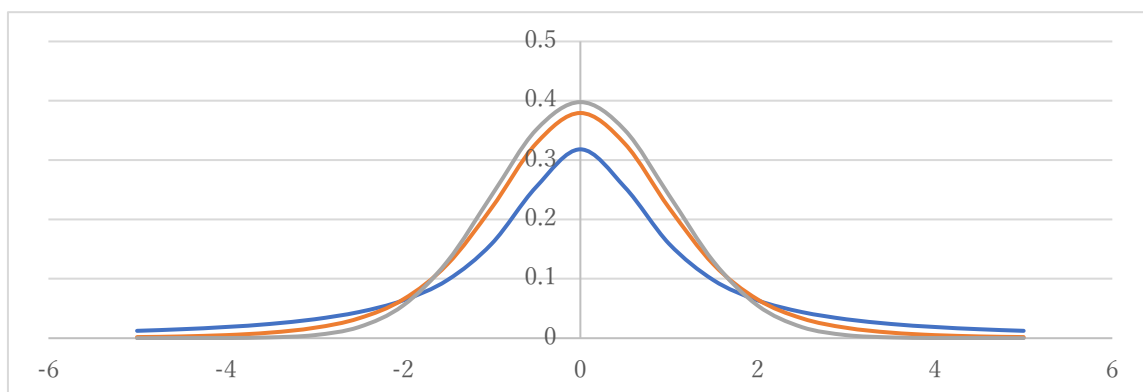
## ・正規分布

言うまでもなく正規分布は統計学において最も重要な分布で、長さや重さを測るときに生じる測定誤差が従う分布として知られる数学モデルです。生物の体長などもこの分布に従うことが知られています。平均から極端に離れたものが殆ど表れないという現象は正規分布で近似できるものが多くあり、品質管理などにも使われます。昔の通知表の5段階評価は、得点の分布が正規分布に従うと仮定して1, 2, 3, 4, 5を割り振っていました。数学的に言えば、二項分布の $n$ を限りなく大きくしたときの極限に現れる分布が正規分布です（ド・モアブルラプラスの定理）。正規分布は平均を山の中心とした左右対称の滑らかな小山型をしています。数表やアプリで簡単に計算できるのは平均0, 分散1の正規分布（標準正規分布）に従う場合の確率で、そのグラフは次のようになります。



## ・t分布 ビジネスや現実社会では分散も平均も分からない状況で十分な情報を得るのが難し

いことが多く、その中で統計的な意思決定をする場合等に用いられるのがt分布です。例えば、20人を選んでレストランA, Bでランチを食べて点数をつけてもらい、A, Bの評価に差があるかどうかを判定するというような顧客満足度調査に使うことができます（対応のある平均値の差の検定）。数学的にはt分布は正規分布から導かれ、その数式は煩雑な形をしています。分布の形は綺麗な小山型をしています。標本数 $n$ が大きいときはほぼ正規分布になりますが、 $n$ が小さいときは山が平べったく裾が遠くまで広がるような形になります。下のグラフは青、オレンジ、グレーの順に $n=2,6,100$ のものです。 $n=100$ の曲線はほぼ正規分布です。



## ・カイ 2 乗分布

カイ 2 乗分布は、おおざっぱに言うと、観測度数と期待度数の食い違い（の指標）が従う分布で、等分散検定、適合度検定、独立性検定などに使われる分布です。例えば新製品の開発の場面で、商品 A、B のどちらが良いか社内でアンケートを取って A と B の評価の優劣を判定するときに使えます（適合度検定）。数学的には、“正規分布に従うデータを標準化した変量を 2 乗して足し合わせたもの”の分布法則がカイ 2 乗分布になります。分布の形状は、標本数  $n$  によって形が変わります。下のグラフは青、オレンジ、グレーの順に  $n=3, 7, 15$  のものです。

