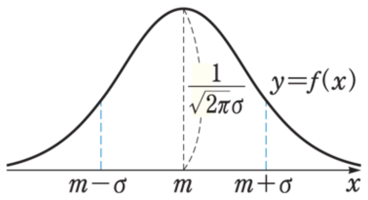
正規分布

身長や一定の規格で作られた部品の誤差など，自然界や社会において偶然に起こる現象には，正規分布と呼ばれる分布に従うものが多い。



mを実数，σを正の実数として，確率変数Xの確率密度関数$f(x)$が，

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-m)^2}{2\sigma^2}}$$

であるとき，このXの確率分布を**正規分布**といい，$\boldsymbol{N(m,\sigma^2)}$で表す。

このとき，Xは正規分布$N(m,\sigma^2)$に従うといい，曲線 $y=f(x)$ を**正規分布曲線**という。

また，次のことが知られている。

正規分布の平均，標準偏差

X が正規分布 $N(m,\ \sigma^2)$ に従う確率変数であるとき，

平均 $E(X)=m$　　　標準偏差 $\sigma(X)=\sigma$

標準化（標準正規分布）

正規分布$N(0,\ 1)$を**標準正規分布**といい，確率密度関数は次のようになる。

$$f(z)=\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}$$

確率変数Zが標準正規分布$N(0,1)$に従うとき，確率 $P(0\leqq Z\leqq u)$の値をuのいろいろな値に対して計算して表にまとめたものを**正規分布表**という。

確率変数Xが正規分布$N(m,\sigma^2)$に従うとき，Xを1次式で変換してできる確率変数$aX+b$も正規分布$N(am+b,a^2\sigma^2)$に従うことが知られている。

$Z=\frac{X-m}{\sigma}$ を考えると，$E(X)=m$，$\sigma(X)=\sigma$ であるから，

$$E(Z)=E\left(\frac{X-m}{\sigma}\right)=\frac{1}{\sigma}E(X)-\frac{m}{\sigma}=0 \qquad \sigma(Z)=\sigma\left(\frac{X-m}{\sigma}\right)=\left|\frac{1}{\sigma}\right|\sigma(X)=1$$

となり，確率変数Zは，標準正規分布$N(0,1)$に従う。このZのことを，Xを**標準化した確率変数**という。

正規分布と標準正規分布

確率変数 X が正規分布 $N(m,\ \sigma^2)$ に従うとき，$Z=\dfrac{X-m}{\sigma}$ とすると，確率変数 Z は標準正規分布 $N(0,\ 1)$ に従う。

確率変数Xが正規分布$N(m,\ \sigma^2)$に従うときも，Xを標準化し標準正規分布$N(0,\ 1)$に変換することにより，Xの確率を求めることができる。

～「標準化」をすることのメリット～

メリット①　どんな正規分布でも，「標準正規分布 $\boldsymbol{N(0,1)}$」に統一でき，標準正規分布1つですべての正規分布の確率が計算できる

メリット②　異なる分布に従うデータの値を，同じ基準で比較することができる

メリット①　どんな正規分布でも，「標準正規分布 $\boldsymbol{N(0,1)}$」に統一でき，標準正規分布1つですべての正規分布の確率が計算できる

問1（文化祭の電子くじ）

ある学校の文化祭では，来場者を 1000 人と見込んでいる。文化祭のあるブースで電子くじを引くと得点X が表示される。X は$N(60,82)$に従うとする。70 点以上が出たら景品を渡すとき，景品はおよそ何個必要か。

メリット②　異なる分布に従うデータの値を，同じ基準で比較することができる

問2（2 科目のテスト比較）

太郎さんはある数学のテストで 70 点，ある英語のテストで 78 点をとった。そのテストの数学の得点X_{math}は$N(60,82)$に従い，英語の得点X_{eng}は$N(70,102)$に従うとする。どちらのテストの方が相対的に高いか，標準化して比較せよ。ただし，どちらのテストも 100 点満点とする。