

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】正規分布の考え方を理解しよう

□正規分布

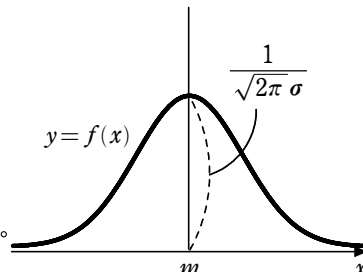
連続型確率変数の分布の代表的なものに、正規分布と呼ばれる分布がある。自然現象や社会現象の中には、観測される値の分布が正規分布に近いものがあり、このとき正規分布が有効に利用される。

m を実数, σ を正の実数とする。このとき, 関数

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

は, 連続型確率変数 X の確率密度関数となることが知られている。

このとき, X は **正規分布 $N(m, \sigma^2)$** に従うという。



$N(m, \sigma^2)$ の N は, 正規分布を意味する英語 normal distribution の頭文字である。

ここで, e は無理数の定数で, $e = 2.71828\dots$ である。曲線 $y = f(x)$ を **正規分布曲線** という。

自然対数の底 e は数学Ⅲの微分法で学びます

正規分布について, 次のことが知られている。

連続型確率変数の期待値, 分散, 標準偏差については, 83 ページで説明している。

正規分布に従う確率変数の期待値, 標準偏差

確率変数 X が正規分布 $N(m, \sigma^2)$ に従うとき

期待値は $E(X) = m,$

標準偏差は $\sigma(X) = \sigma$

N (期待値, 分散)

正規分布はアブラム・ド・モアブルによって1733年に導入された。

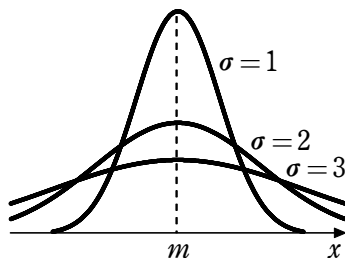
その後、1809年のカール・フリードリヒ・ガウスによる誤差論で詳細に論じられたため、ガウス分布とも言われる。

確率変数 X が正規分布 $N(m, \sigma^2)$ に従うとき,

X の分布曲線 $y = f(x)$ は, 次のような性質をもつ。

正規分布曲線の性質

- 1 直線 $x = m$ に関して対称であり, y は $x = m$ で最大値をとる。
- 2 x 軸を漸近線とし, x 軸と分布曲線の間の面積は 1 である。
- 3 標準偏差 σ が大きくなると曲線の山は低くなって横に広がる。 σ が小さくなると曲線の山は高くなって, 直線 $x = m$ の周りに集まる。



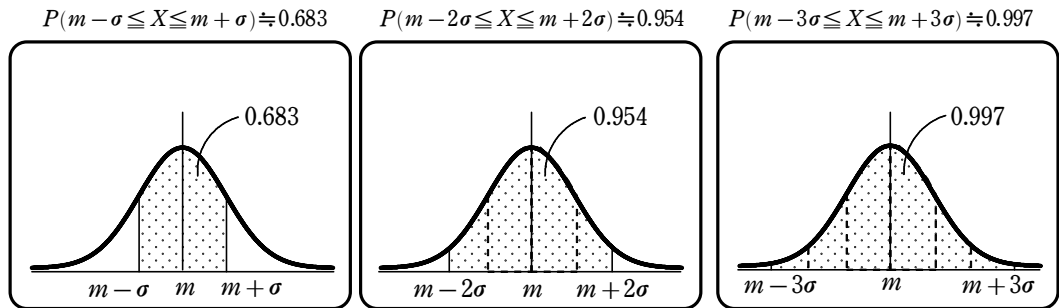
標準偏差が大きくなる⇒散らばる
⇒なだらかになる
標準偏差が小さくなる⇒集中する
⇒山が高くなる

<補足> 曲線が限りなく近づくと直線とその曲線の漸近線という。



正規分布曲線

また、確率変数 X が正規分布 $N(m, \sigma^2)$ に従うとき、その確率について次が成り立つことが知られている。



このことは、たとえば X が $m - \sigma \leq X \leq m + \sigma$ を満たす確率は約 68.3 %、
 $m - 2\sigma \leq X \leq m + 2\sigma$ を満たす確率は約 95.4 %であることを意味する。

コラム 偏差値

試験の成績を「偏差値」という数値で表すことがあります。平均点が m 、標準偏差が σ である試験において、得点が X 点であるとき、偏差値 T は以下の式で計算することができます。

$$T = 10 \times \frac{X - m}{\sigma} + 50 \quad \dots\dots ①$$

ここでは、試験の得点 X を確率変数としたとき、 X は正規分布に従うとみなして考えてみましょう。得点 X が平均 m 、標準偏差 σ の正規分布に従うとみなすと、

① の式で得られる偏差値 T の平均（期待値）は $10 \times \frac{m - m}{\sigma} + 50 = 50$ 、

標準偏差は $\left| \frac{10}{\sigma} \right| \times \sigma = 10$

上の結果から、偏差値 T は平均が 50、標準偏差が 10 である正規分布に従うとみなすことができます。

このとき、 $P(50 - 2 \cdot 10 \leq T \leq 50 + 2 \cdot 10) \approx 0.954$
 すなわち $P(30 \leq T \leq 70) \approx 0.954$

となり、偏差値が 30 から 70 の間に受験者の約 95.4 % が含まれることがわかります。

したがって、偏差値 70 以上の受験者は全体の約 2.3 % であることから、受験者が 1 万人であれば、偏差値 70 である人は大体 230 位の位置にいると考えられます。

補足

自然界の事象の中には正規分布に従うものがあることが知られています。また、そのままでは変数が正規分布に従わない場合もその対数をとると正規分布に従う場合があります。

正規分布が統計学上特別な地位を持つのは中心極限定理が存在するためです (p.92 で出てきます)。自然界の事象の中には、正規分布に従うものがあることが知られていますが、それは必ずしも多数派というわけではありません。19 世紀ではさながら「正規分布万能主義」といったものがまかり通っていましたが、20 世紀以降そういった考え方に修正が見られ、今日においては社会現象、生物集団の現象等々、種別から言えば、正規分布に従うものはむしろ少数派であることが確認されています。