

# データサイエンス6th

『正規分布』

# 今日のテーマ

○正規分布の基本を知ろう！

どんな形？

どんな特徴？

○身の回りにある正規分布を探そう！

どんなデータに当てはまる？

○平均とばらつきがグラフにどう影響するか見よう！

グラフの形が変わるってどういうこと？

○データ分析でなぜ正規分布が大切なのか理解しよう！

# 本時の目標

- ・正規分布の概念を理解し、その特徴を説明できる。
- ・身近なデータが正規分布に従う例を知り、データ分析における正規分布の重要性を認識する。
- ・正規分布と標準偏差の関係を理解し、データのばらつきを考察できる。

# データってどんな形をしている？

- ・ みんなのテストの点数をグラフにしたらどうなる？
- ・ クラスの身長をグラフにしたらどうなる？

# 正規分布の「形」

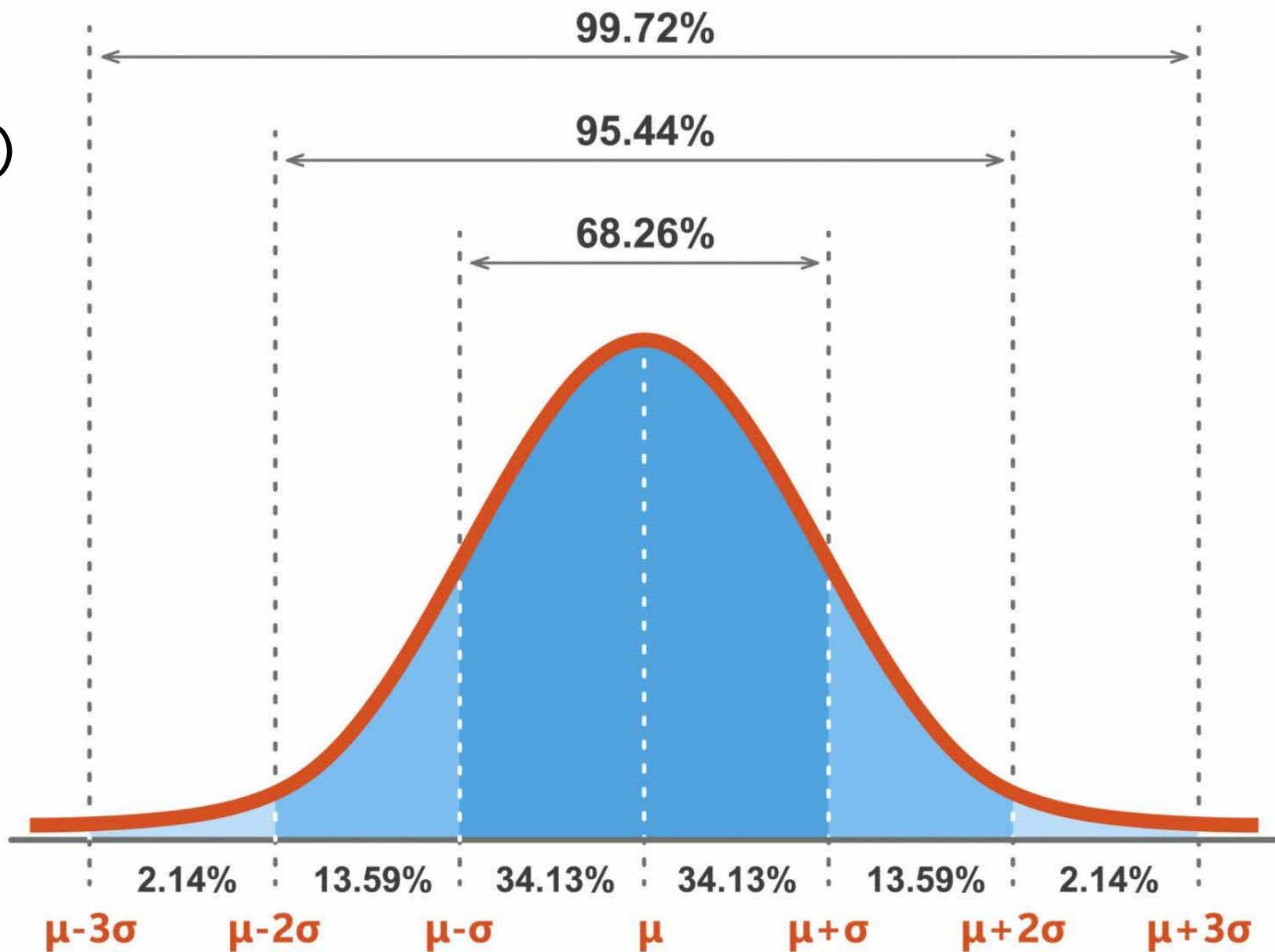
多くのデータが平均値の近くに集まっていて、そこから離れるにつれてデータの数が減っていく形。

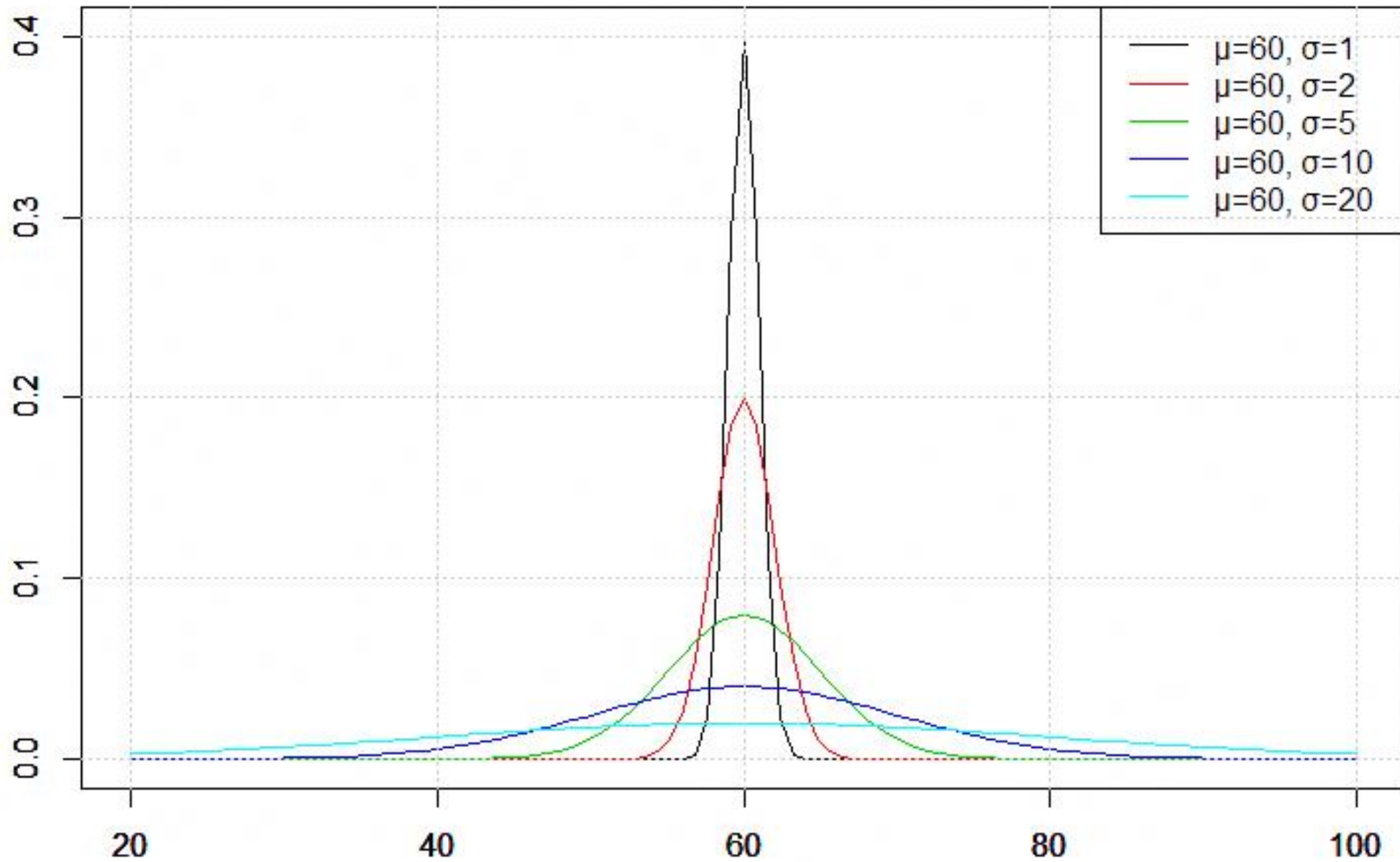
まるで「釣鐘」（ベルカーブ）のような形をしているのが特徴です。

自然界や社会で起こる多くの事象が、この形に近いことが知られている。「ごく普通のこと」が最も多く、平均から離れるほど少なくなる。

平均值 ( $\mu$ )

標準偏差 ( $\sigma$ )





<https://bellcurve.jp/statistics/course/7797.html?srsId=AfmBOor64nANq3m9wMcvCLUIV-vz0QicUK7zaSCxoRE1KmgIv2W68ZjF>

**特徴1：**

**左右対称** グラフの真ん中を境に、左右が鏡のように同じ形をしています。

**特徴2：**

**中央に集中** 一番高いところがデータの真ん中（平均値、中央値、最頻値）になります。

# 平均値と正規分布

平均値 ( $\mu$ ) : データの中心を表す値。  
正規分布のグラフでは、この平均値がグラフの「てっぺん」の位置になります。

見てみよう！  
平均値が変わると、グラフの形は変わらず、左右に移動します。

# 標準偏差と正規分布

標準偏差 ( $\sigma$ ): データの「ばらつき」や「散らばり」の度合いを表す値。

- ・ 標準偏差が大きい  $\rightarrow$  データが広く散らばっている (グラフが平べったい)
- ・ 標準偏差が小さい  $\rightarrow$  データが狭い範囲に集中している (グラフがとがっている)

見てみよう!

標準偏差が変わると、グラフの広がり (とんがり具合) が変わります。

# ワーク「身の回りにおける正規分布の例を探す」

実は、私たちの身の回りには正規分布に従うデータがたくさんあります。そのようなデータを探してみましょう。また、スライドにまとめてみましょう。

どのようにデザインするかは各自に委ねます。ただし、枚数は1枚とします。

4つのカテゴリーに分けます。以下から選んでください。

①自然・身体、②社会・教育、③工業製品・品質管理、④金融・ビジネス

スライドが完成したらグループで共有します。

# 身の回りにおける正規分布の例

## 人の身長や体重

日本人の男性の身長は、平均値の170cm台に多くの人が集集中し、それより高い人や低い人は少なくなります。

## テストの点数

クラスのテストの点数も、平均点付近に多くの生徒が集まり、高得点や低得点の生徒は少数派になることが多いです。

## 製品の品質

工場で作られる製品の重さや長さなども、目標値を中心にばらつきが生じ、正規分布になる傾向があります。

# 正規分布がなぜ大切なのか？

データを扱う上で、正規分布を理解することはとても重要です。

## ①データの傾向を素早く把握できる

正規分布の形をしているデータは、平均値と標準偏差を見るだけで、全体の傾向やばらつき具合がわかります。

## ②異常なデータを見つけやすい

正規分布から大きく外れたデータは、何か特別な要因がある「異常値」かもしれないと判断できます。

## ③AIやデータサイエンスの基礎

今後、AIや機械学習でデータを分析する際に、正規分布の知識は様々な統計的手法やモデルの理解に役立ちます。

# まとめ

- ・正規分布は、**左右対称の釣鐘型**のグラフで表されるデータの分布です。
- ・**身長**や**テストの点数**など、身の回りの様々なデータに現れます。
- ・**平均値**はグラフの**中心**を、**標準偏差**はグラフの**広がり**を示します。
- ・正規分布の理解は、**データ分析の第一歩**としてとても大切です。