心理統計法第 04 回　　　　　　　　　学籍番号：
講義ノート p.1
講義担当：佐々木康成　　　　　　　名　　前：

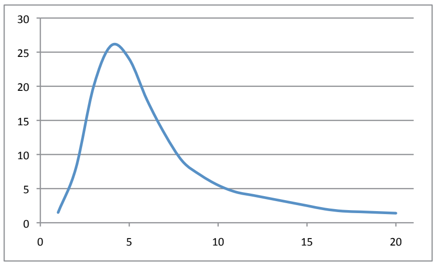
【手がかり用スペース】

ばらつき具合の視覚化
- 手にしたデータを整理して直感的に理解す るには？
  - 表にする
    - 度数分布表
  - 図にする
    - ヒストグラム・度数多角形

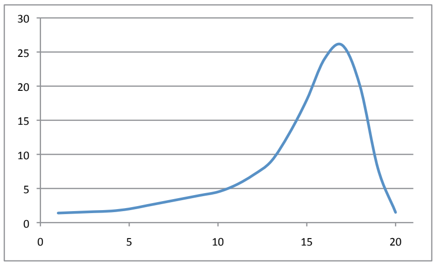
度数分布表
- 級間を決定
  - 真の限界（Real Limits）
  - 中心点（＝階級値）
- 度数をカウント
- 表の形で表現する

ヒストグラム・度数多角形
- 度数分布表を元に作成
  - ヒストグラム
    - 棒グラフ状にデータを表記
  - 度数多角形
    - 中心点を結び、多角形状にデータを表記
- ともにより直感的な理解が可能

代表値
- 集団の性質を一つの数値で表現する
  - 平均値
  - 中央値
  - 最頻値
- どの代表値にするかは、集団の性質を見極 めた上で決定
- 平均値と中央値で分布の「歪み」が決まる



正に歪んだ分布の例



負に歪んだ分布の例

【要点書き出し用スペース】

学籍番号：

名　前：

【手がかり用スペース】

追補情報

歪度と尖度

左右対称かどうかは「歪度」で計算できる

どれだけ尖っているかは「尖度」で計算で きる

練習問題の解答

問 1：最小値・最大値・範囲を求めなさい

最小値：23

最大値：83

範囲：60

問 2：このデータをもとに度数分布表を作成しなさい

問 3：このデータをもとにヒストグラムを作成しなさい

問 3：このデータにおいて、平均値・中央 値・最頻値のそれぞれの値を答えなさい

平均値：48.8

中央値：45.0

最頻値：度数分布表の級間の設定によって変化
二つ出た場合は二つ書くこと。

問 4：このデータは正に歪んでいるといえる か、負に歪んでいるといえるか、答えなさい。

正に歪んでいます。根拠は「中央値＜平均値」。

本日の内容

ばらつき具合とその指標

散布度の求め方

Z スコアと偏差値

使用するデータの特徴

代表値

代表値として「平均値」を用いる

数学的に最も高度な代表値

分布の全ての値を使って求める

分布の形

左右対称

中央部分が最も値が多く、極端な値は 少ない釣り鐘型

「正規分布」

【要点書き出し用スペース】

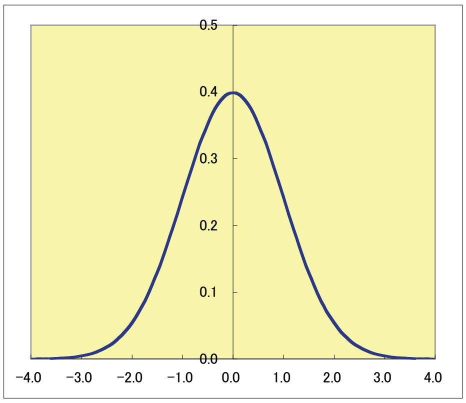
心理統計法第 04 回　　学籍番号：
講義ノート p.3
講義担当：佐々木康成　　名　　前：

【手がかり用スペース】

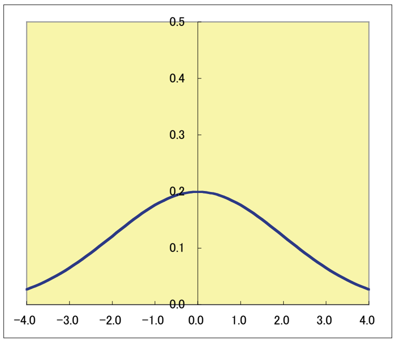
正規分布

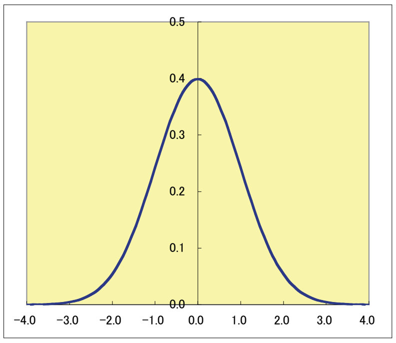
平均値が等しい正規分布の比較
　　平均値が等しいからといっても
　　同じ正規分布ではない

平均値が等しい正規分布の比較
　　下の二つの図は何が違うのか？



正規分布の例





ばらつき具合の違いの表現
　　なだらかな分布
　　　　データがばらついている
　　尖った分布
　　　　データがばらついていない

ばらつき具合と分布の形
　　データのばらつきが分布の形を決める

ばらつき具合を他人に伝えるためには？
　　数値で表現する
　　「散布度」という指標を用いる
　　「代表値」と「散布度」で分布の特徴を表現

【要点書き出し用スペース】

心理統計法第 04 回　　　　　　　　　学籍番号：
講義ノート p.4
講義担当：佐々木康成　　　　　　　　名　　前：

【手がかり用スペース】

簡単な散布度

範囲

最大値と最小値から求める値

極端な値に弱い

2 つの値しか使っていないので、情報量が少ない

平均値のように「全てのデータ」を使いたい

ばらつき具合の表現

ばらつきとは「何からの」ばらつきか？

代表値（平均値）からどれだけ離れているか

それぞれの値が平均値からどれだけ離れているかを計算して

集計すれば良い

個々の値がどれだけ平均値から離れているか以下の空欄を埋めること

散布度の計算過程			
名前	得点	得点−平均値	(得点−平均値)^2
A	2		
B	2		
C	3		
D	3		
E	5		
F	6		
G	6		
H	7		
I	8		
J	8		
	平均値	合計値	合計値 = 偏差平方和
	5		
		分散 = 偏差平方和/データ数	
		標準偏差 = √分散	

【要点書き出し用スペース】

心理統計法第 04 回　　　　　　　　**学籍番号：**
講義ノート p.5
講義担当：佐々木康成　　　　　　　**名　　前：**

【手がかり用スペース】

偏差
代表値からどれだけ離れているか
偏差 ＝ 個々の値 - 平均値
p.4 の表に書き入れましょう

偏差の合計
偏差を出し終えたら合計する
0 になる
証明は以下の通り

偏差の平均
= 偏差の合計 / データ数
=（(それぞれのデータから平均値を引いたもの）の合計）÷ データ数
=（全データの合計から平均値 × データ数を引いたもの）÷ データ数
=（全データの合計 ÷ データ数）ー（平均値 × データ数 ÷ データ数）
= 平均値ー平均値
=0

従ってそのままでは立ち行かなくなる

偏差二乗和
なぜ偏差の合計は 0 になるかは上の証明通り
ならば全てを正の数にすればいい
二乗すれば正の数になる（単位も二乗になる）
全て計算したら合計する
p.4 の表に書き入れましょう

偏差二乗和（Sum of Squares = SS）
「偏差二乗和」
「偏差自乗和」
「偏差平方和」
どれも同じものを指す
最も基本となる散布度の指標
データ数が増えると値が大きくなる

【要点書き出し用スペース】

心理統計法第 04 回　　　　　　　　　　　学籍番号：
講義ノート p.6
講義担当：佐々木康成　　　　　　　　　　名　　前：

【手がかり用スペース】

分散
　偏差二乗和をデータ数で調整した値
　　分散＝偏差二乗和 ÷ データ数
　　調整してあるので、ばらつき具合の評価が比較しやすい
　　ただし単位は二乗のまま

標準偏差
　分散の単位を元の単位に戻したもの
　　分散の平方根（ルート）を取る
　　標準偏差 ＝ √分散
　標準偏差は偏差の二乗平均
　　二乗平均とは、平均したい値を二乗して合計してデータ数で割って√を取ったもの
　p.4 のサンプルデータの標準偏差
　　√ 5.00 ≒ 2.24
　　「このテストの結果は平均 5.0 点で、おおよそのばらつき具合は平均点を中心にしてを 2.24 点である」といえる
　　　もっと詳しくは今回の最後で触れる

散布度についてのまとめ
　代表値と散布度でデータのおおまかな性質を捉えられる
　　代表値
　　　平均値・中央値・最頻値
　　散布度
　　　偏差二乗和・分散・標準偏差

【偏差平方和と分散、標準偏差の関係】
「偏差二乗和」はデータ数によって値が左右される（一般的にデータ数が多い方が値が大きくなる)。従って、データ数によって調整が必要となる。データ数で割った値を「分散」と呼び、一般的な散布度の指標とする。しかし、分散は偏差平方を求めた際に偏差を二乗した値であるため、元の単位とは隔たりがある。そこで、単位を戻すために分散の根を取る。これが「標準偏差」である。

【要点書き出し用スペース】

心理統計法第 04 回　　　　　　　　学籍番号：
講義ノート p.7
講義担当：佐々木康成　　　　　　　名　　前：

【手がかり用スペース】

データを比較する

「異なる科目のテストで同じ点数を取ったとする。それぞれの科目の平均値が同じなら成績は等しくなるか？」

実は散布度を考慮していない設問

「各科目の成績はどれくらいばらついているか」の情報が必要

平均 50 点のテストで 60 点を取った場合

標準偏差 10 点の分布

60 点は標準偏差 1 つ分上回っているといえる

標準偏差 5 点の（10 点よりもばらつきが少ない）分布

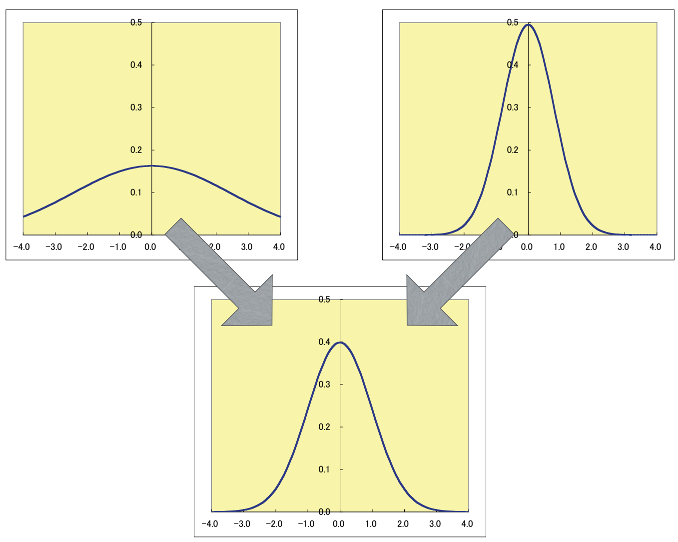
60 点は標準偏差 2 つ分上回っているといえる

標準偏差を単位として個々の得点がどれだけ平均値から離れているかが数値化できる

偏差 ÷ 標準偏差＝標準得点

データを全て標準得点に直した分布を「標準正規分布」と呼ぶ

標準正規分布の形に変換することを「標準化」という



Z スコア

（個々の得点 ― 平均値）÷ 標準偏差＝ Z スコア（標準得点）

標準正規分布の形に変換すればどんなデータでも比較可能

異なる科目間

身長と体重

【要点書き出し用スペース】

心理統計法第 04 回　　　　　　　　学籍番号：
講義ノート p.8
講義担当：佐々木康成　　　　　　　名　　前：

【手がかり用スペース】

データの標準化（Zスコア）と偏差値

名前	得点	偏差	Zスコア（偏差/標準偏差）
A	2		
B	2		
C	3		
D	3		
E	5		
F	6		
G	6		
H	7		
I	8		
J	8		

Z スコアの意味

E くん（平均値と完全に一致した成績）

$0 \div 2.24 = 0$

F くん（平均値 +1.0SD の成績）

$1 \div 2.24 \fallingdotseq 0.45$

データの個々の位置を表すのに役立つ

偏差値

Z スコアには小数点以下の値もマイナスの値もある

Z スコアを 10 倍して 50 を足した値が「偏差値」

偏差値 60 は、標準偏差 1 つ分プラスである、という意味

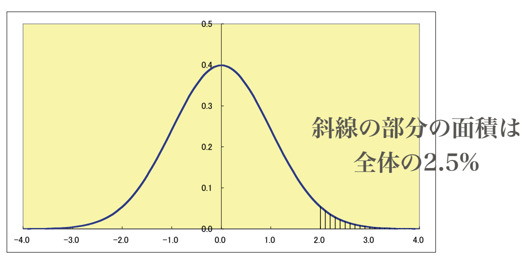
上側確率

標準正規分布の特徴

全体の面積を 1 とした際の

「ある値より上」の面積が

既に計算済み



斜線の部分の面積は全体の 2.5%

標準偏差 ±1 の範囲に、

全体の 68.26%が入る

「このテストの結果は平均 5.0 点で、平均点を

中心とした ±2.24 点の範囲に全体の 68.26%

のデータが入る」といえる

【要点書き出し用スペース】