

# データサイエンス概論

## 第4回 第1講

### 事前予習

インターネットで、以下のようなキーワードで検索を行い大まかな内容に目を通しておく。

- 標準偏差
- 正規分布

授業で基本的なことが説明されるので、そこに出てくる用語などを確認しておくだけで良い。

注：

スライド：1-10

分散を、 $N$ で割るのか  $N-1$  で割るのか、という点については長い議論がある。ここでは  $N-1$  で割るという多くの教科書で採用されている方にしておく。

# データサイエンス概論

## 第4回 第2講

### 事前予習

インターネットで、以下のようなキーワードで検索を行い大まかな内容に目を通しておく。

- 統計 仮設検定
- 独立変数、 従属変数
- 相関係数

注：独立変数の概念について調べるのは、授業の説明を一度聞いてからでも構わない。何が独立であるのか。基本的には他の要素によって影響を受けて変化しない、ということであり、影響を受けて変化するなら、独立ではない、ということになる。

# データサイエンス概論

## 第4回3講

### 事前予習

インターネットで、以下のようなキーワードで検索を行い大まかな内容に目を通しておく。

- t 検定
- カイ二乗検定
- ANOVA (分散分析)

# データサイエンス概論

## 第4回 4講

### 事前予習

インターネットで、以下のようなキーワードで検索を行い大まかな内容に目を通しておく。

- ANOVA（分散分析） 一元配置 二元配置

データサイエンスを業務として行う場合にどのようにプロセスを進めていくかということについては多くの説明や提案があることを理解しておくこと。インターネットで調べるとビジネス関係の応用が多く出てくる。授業ではアカデミックな応用も念頭に置いたものに修正したものを提示している。

注：

4-11: 修正： 「二次元配置」 → 「二元配置」