

データの見方

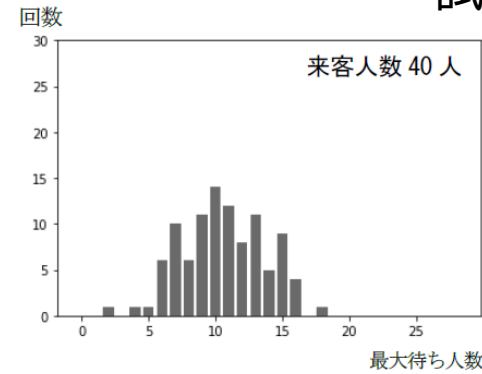
この時間の目標

- データの見方
- S データの様々な見方がよく理解でき、配慮して活用しようと思った
- A データの様々な見方がよく理解できた
- B データの様々な見方が理解できた
- C データの様々な見方が理解できなかった

グラフ形状の変化

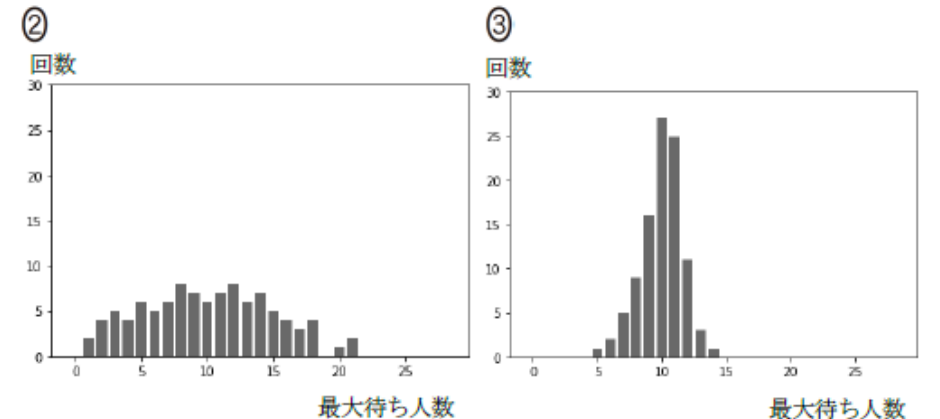
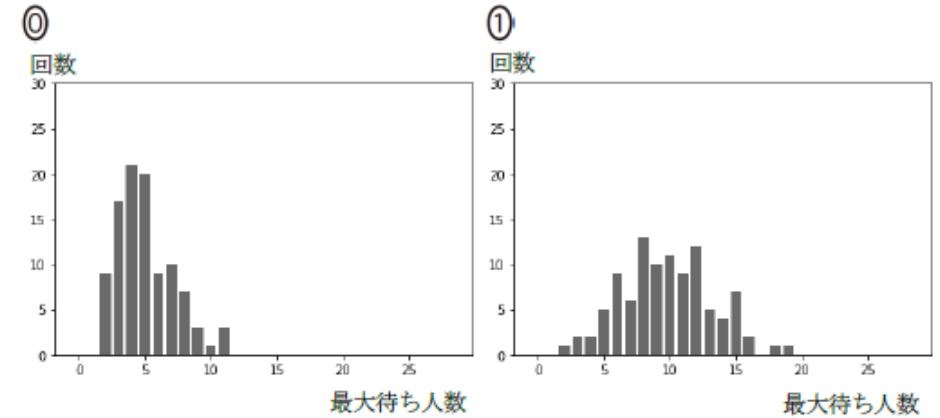
グラフ形状の変化

- 来客40名・処理時間4分の
窓口シミュレーション結果



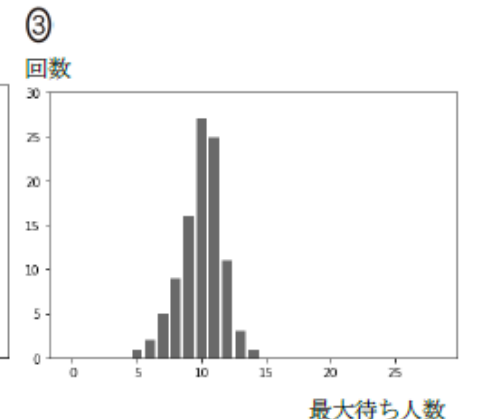
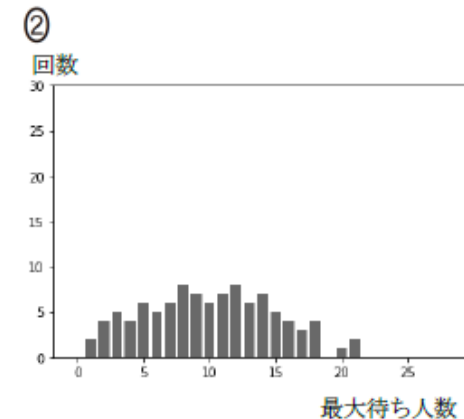
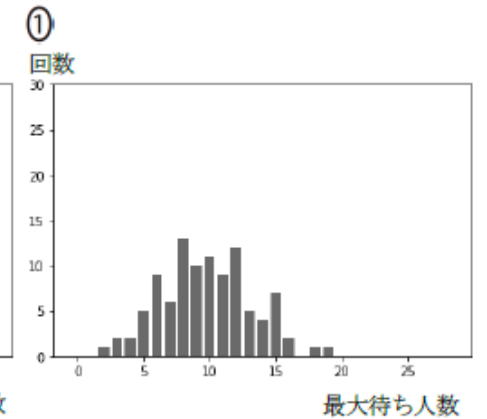
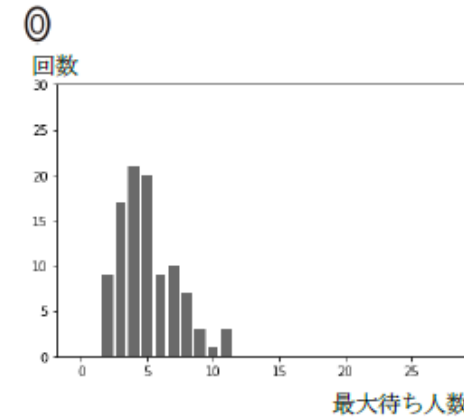
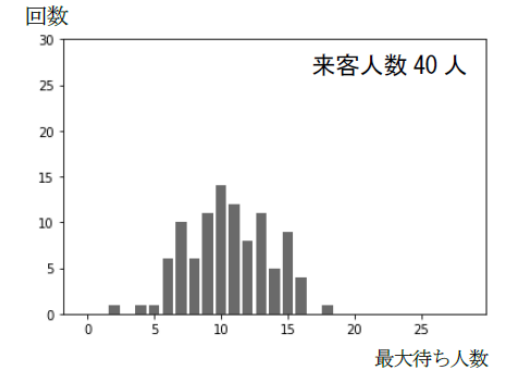
ペアで考える

- 来客40名・処理時間3分の
窓口シミュレーション結果は
どうなる？
右の0～3のうちから選べ



グラフ形状の変化

- 来客40名・処理時間4分のシミュレーション
- 来客40名・処理時間3分なら



実際にシミュレーションすると

グラフ形状の変化

- 状況が変わるとどう変化するか
 - データが変わるとどう変化するか
 - 論理的に理由があるはず
-
- いろいろなグラフで「考える」トレーニングを

散布図行列

散布図行列

- 図1の I ~ IV は、それぞれの項目の全参加チームのヒストグラムを決勝進出チームと予選敗退チームとで色分けしたものであり、① ~ ④ は決勝進出チームと予選敗退チームに分けて作成したヒストグラムである。あ ~ か は、それぞれの二つの項目の全参加チームと決勝進出チーム、予選敗退チームのそれぞれに限定した相関係数である。また A ~ F は、それぞれの二つの項目の散布図を決勝進出チームと予選敗退チームをマークで区別して描いている。例えば、図1の A は縦軸を「1試合当たりの得点」、横軸を「1試合当たりのショートパス本数」とした散布図であり、それに対応した相関係数はあで表されている。

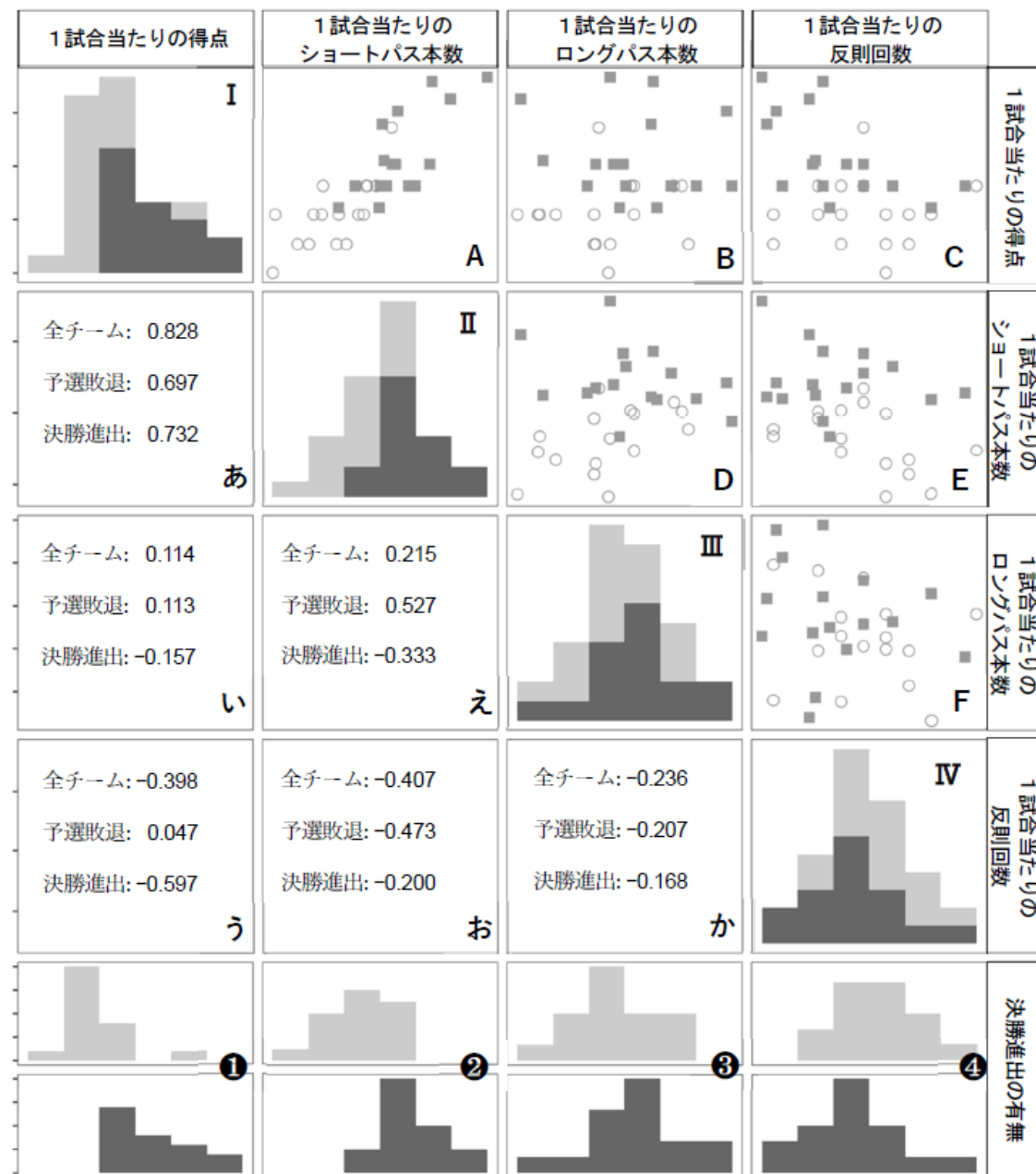


図1 各項目間の関係

問1 次の問いに答えよ。

- a 次の文章を読み, 空欄 ア ~ ウ に入れる最も適当なものをそれぞれの解答群のうちから一つずつ選べ。ただし, 空欄 ア・イ の順序は問わない。
 - 図1を見ると, 予選敗退チームにおいてはほとんど相関がないが, 決勝進出チームについて負の相関がある項目の組合せは, 1試合当たりの ア と イ である。また, 決勝進出チームと予選敗退チームとで, 相関係数の符号が逆符号であり, その差が最も大きくなっている関係を表している散布図は ウ である。したがって, 散布図の二つの記号のどちらが決勝進出チームを表しているかが分かった。
- ア・イの解答群
 - 0:得点 1:ショートパス本数 2:ロングパス本数 3:反則回数
- ウの解答群
 - 0:A 1:B 2:C 3:D 4:E 5:F

問1 次の問いに答えよ。

- b 図1から読み取れることとして誤っているものを解答群から一つ選べ。工
- 工 の解答群
 - 0:それぞれの散布図の中で、決勝進出チームは黒い四角形(■), 予選敗退チームは白い円(○)で表されている。
 - 1:全参加チームを対象としてみたとき, 最も強い相関がある項目の組合せは1試合あたりの得点と1試合あたりのショートパス本数である。
 - 2:全参加チームについて正の相関がある項目の組合せの中には, 決勝進出チーム, 予選敗退チームのいずれも負の相関となっているものがある。
 - 3:1試合あたりのショートパス本数の分布を表すグラフ②で, 下の段は決勝進出チームのヒストグラムである。

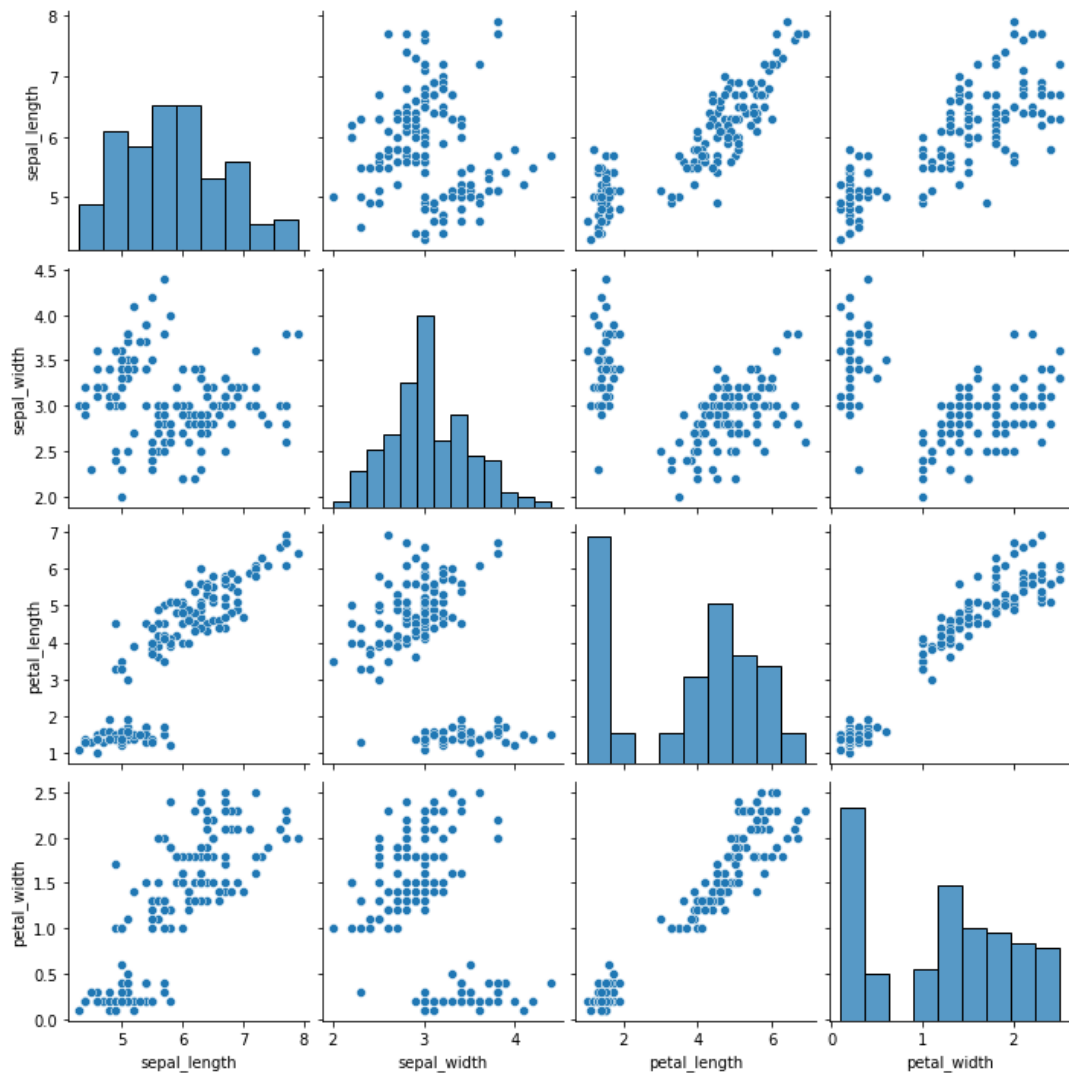
散布図行列

- 図の説明は書いてある
 - ゆっくり読んで考えればできる
- ゆっくり読むと時間が足りない
 - 散布図・ヒストグラム・相関係数 を理解しておく

散布図行列

```
import matplotlib.pyplot  
import seaborn
```

```
iris = seaborn.load_dataset( "iris" )  
seaborn.pairplot( iris )  
matplotlib.pyplot.show()
```



増えてるの？減ってるの？

データをわかりやすく表現

- 2022年4月-5月の東京都における感染者数
 - サンプルデータ1 covid.xlsx
- 感染者が増加傾向？減少傾向？
- わかりやすく伝えるには

昇順と降順

番外編

JupyterNotebookを準備

繰り返しを使って

- 1から10を順に表示するプログラム

- 10から1を順に表示するプログラム