

レポート提出票

科目名: 情報工学実験1

実験課題名: 課題4 統計学入門

実施日: 2020年 8月 3日

学籍番号: 4619055

氏名: 辰川力駆

共同実験者:

1 実験の目的

(1) 離散分布の理解

代表的な離散分布である二項分布を理解する。

(2) 標本平均の分布

中心極限定理を理解する。

(3) 連続分布の理解

代表的な連続分布である正規分布を理解する。

2 実験 1

二項分布を基にした実験

2.1 目標

離散分布の基本として二項分布の特徴を理解する。

2.2 実験手順

(1) “赤 4 個、白 6 個の合計 10 個の玉が入った箱から 1 つ取り出して色を確認後、取り出した玉を箱に戻す作業を 5 回繰り返す。” というのを 1 回の試行として、これを 30 回試行して記録する。

(2) 取り出された赤玉の回数の度数表と棒グラフを Excel で作成する。

(3) 取り出された赤玉の個数の平均値を求める。

2.3 実験結果

以下に実験結果の表とグラフを示す。また、取り出された赤玉の個数の平均値は約 2.433333333 となった。

表 1: 試行における赤玉の個数

試行	回数					赤玉の個数
	1	2	3	4	5	
1	○	○	○	×	○	4
2	×	○	×	○	×	2
3	○	×	○	○	×	3
4	×	○	×	×	○	2
5	×	×	○	○	×	2
6	○	×	○	×	×	2
7	○	×	○	×	○	3
8	○	×	○	○	○	4
9	×	×	×	○	×	1
10	×	×	○	×	×	1
11	×	○	×	×	○	2
12	×	○	○	○	×	3
13	○	○	○	○	×	4
14	○	○	×	○	×	3
15	○	○	○	×	×	3
16	○	×	○	○	○	4
17	×	×	×	×	×	0
18	×	×	×	○	○	2
19	○	×	×	○	○	3
20	○	×	○	○	×	3
21	×	○	○	×	○	3
22	○	×	×	×	×	1
23	×	×	×	×	○	1
24	×	○	○	○	○	4
25	×	×	○	○	○	3
26	○	×	×	×	○	2
27	○	×	○	×	×	2
28	○	○	×	○	○	4
29	○	×	×	×	×	1
30	×	×	○	×	×	1
平均						2.43333333

表 2: 出現回数の度数表

赤玉の個数	度数
0	1
1	6
2	8
3	9
4	6
5	0

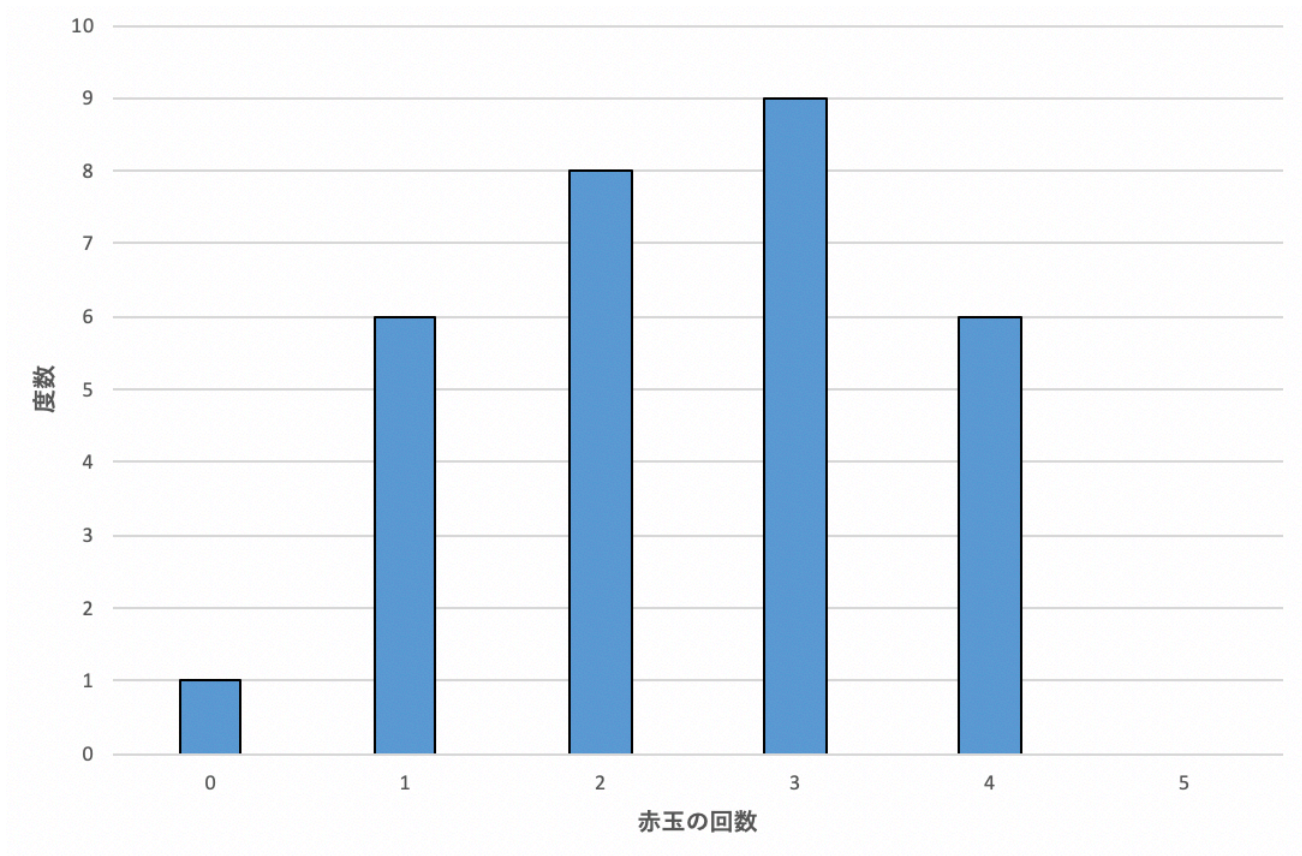


図 1: 赤玉出現回数の棒グラフ

2.4 レポート課題

レポート課題 1-1

実験 1 の結果をまとめよ。

表 2 の度数分布表より, 最頻値は 1 であり, 表 1 の実験結果より, 平均値は 1.83 である. 中央値は 2 である. ヒストグラムから見ると, 1 の頻度が結構高いので, 平均値が中央値より小さいと考えられる.

レポート課題 1-2

実験 1 の結果は、二項分布に従っているといえるか述べてよ。

実験 1 の結果は、二項分布に従っていない。

3 実験 2

中心極限定理を基にした実験

3.1 目標

- データを視覚的に表現する方法を理解する
- データの正しい要約を行う

3.2 実験手順

- (1) 年収データの平均値・中央値・最頻値を求める
- (2) 年収を 50 万円刻みで区切ったヒストグラムを作成する
- (3) 箱ひげ図を作成する

3.3 実験結果

3.4 レポート課題

レポート課題 2-1

実験 2 の結果をまとめよ。

レポート課題 2-2

大数の法則とは何か、実験 2 の結果に基づいて説明せよ。

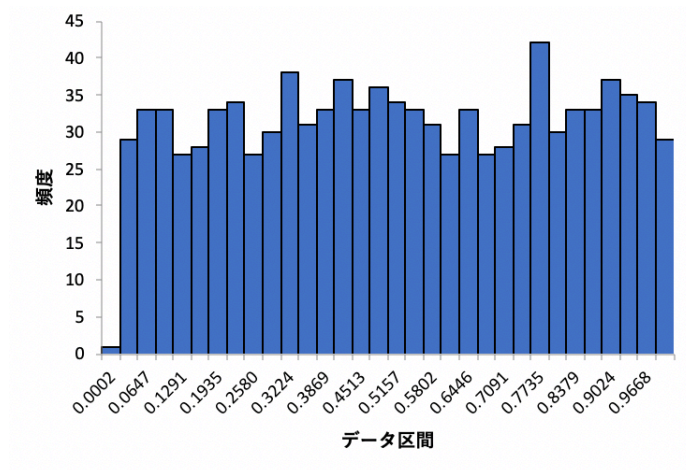


図 2:

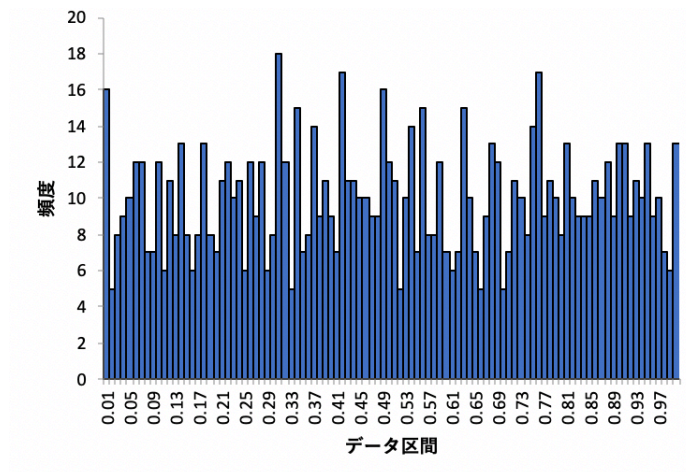


図 3:

レポート課題 2-3

中心極限定理とは何か、実験 2 の結果に基づいて説明せよ。

レポート課題 2-4

区間 $[0,1]$ の一様分布から 12 個選び、その合計から 6 を引いた値を Y とする。 Y はどのような分布に従うか述べよ。

4 実験 3

正規分布を基にした実験

4.1 目標

- データを視覚的に表現する方法を理解する

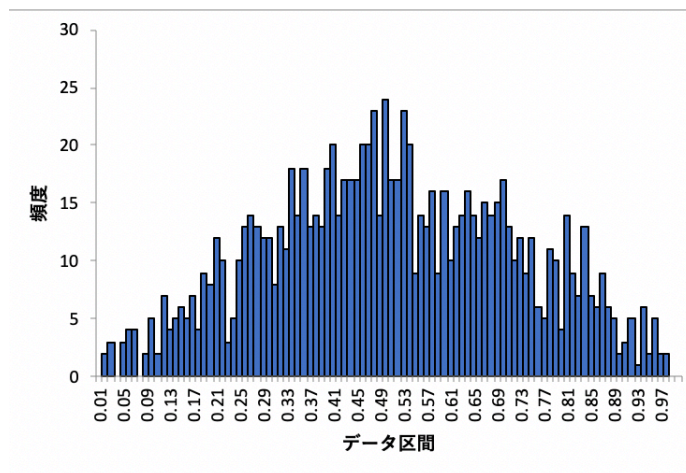


図 4:

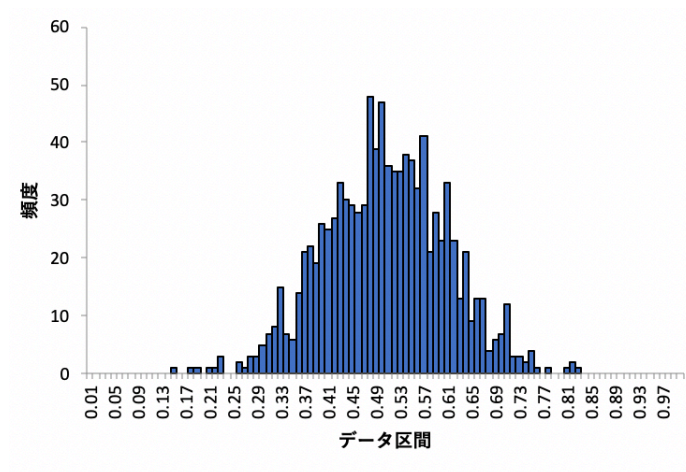


図 5:

- データの正しい要約を行う

4.2 実験手順

- (1) 年収データの平均値・中央値・最頻値を求める
- (2) 年収を 50 万円刻みで区切ったヒストグラムを作成する
- (3) 箱ひげ図を作成する

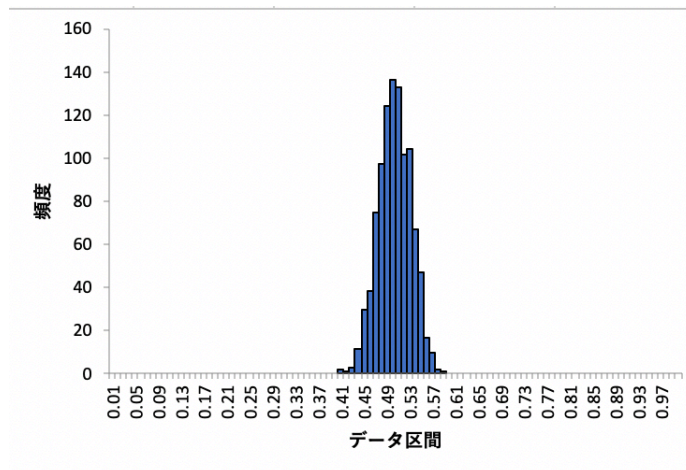


図 6:

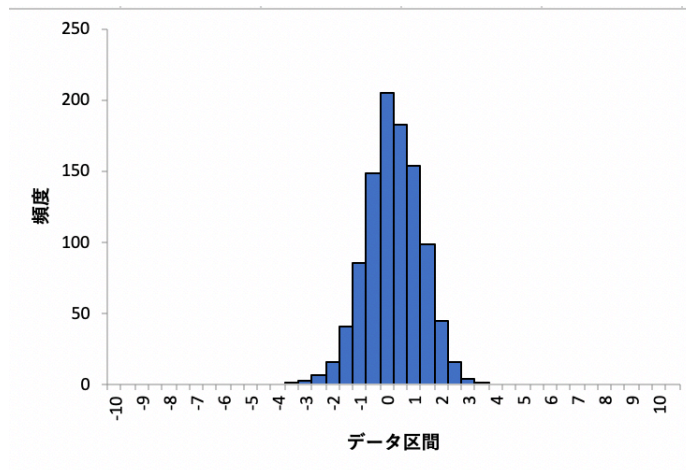


図 7:

4.3 実験結果

4.4 レポート課題

レポート課題 3-1

実験 3-5、3-6 で作成したヒストグラムを比較して、考察せよ。

レポート課題 3-2

実験 3-5、3-7 で作成したヒストグラムを比較して、考察せよ。

レポート課題 3-3

課題 3-1、3-2 から、正規分布における期待値と標準偏差は確率分布の何を決定しているといえるか述べて。

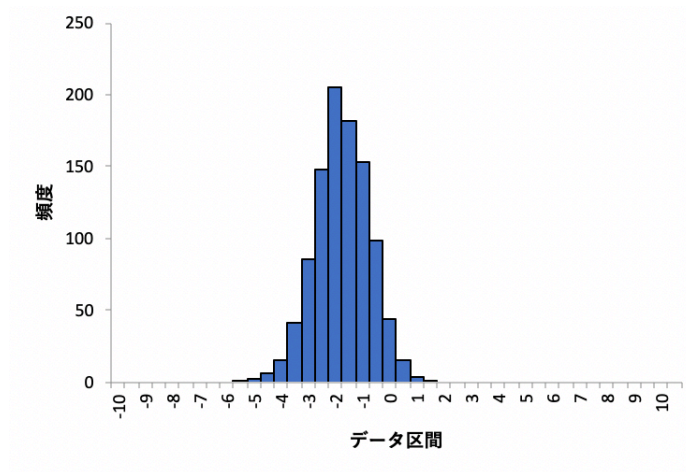


図 8:

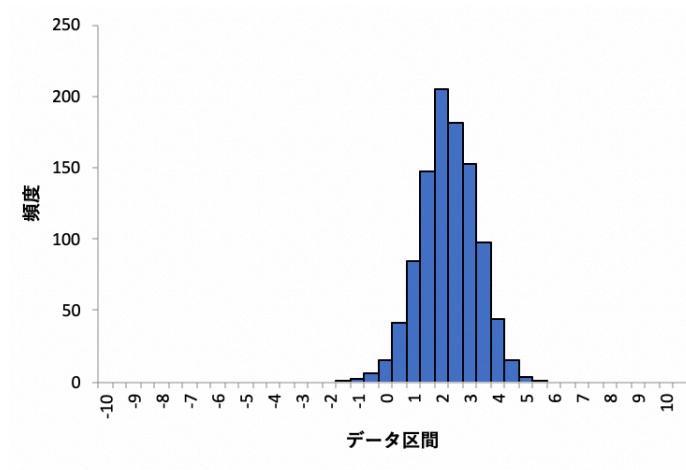


図 9:

レポート課題 3-4

偏差値の確率密度関数を書け。

レポート課題 3-5

試験の点数や順位と比較して偏差値の短所・長所を述べよ。

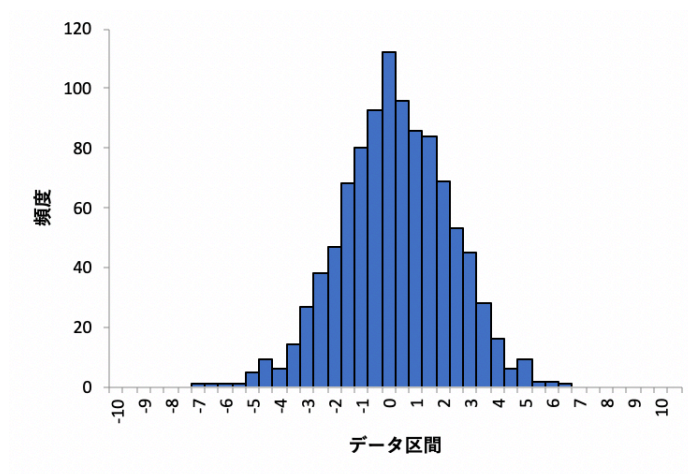


図 10:

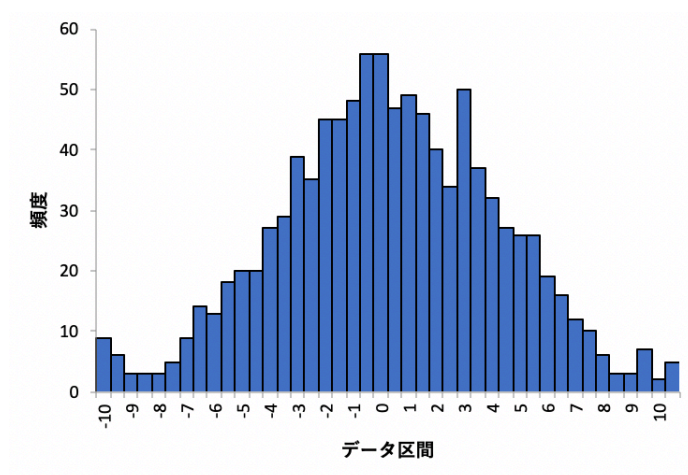


図 11:

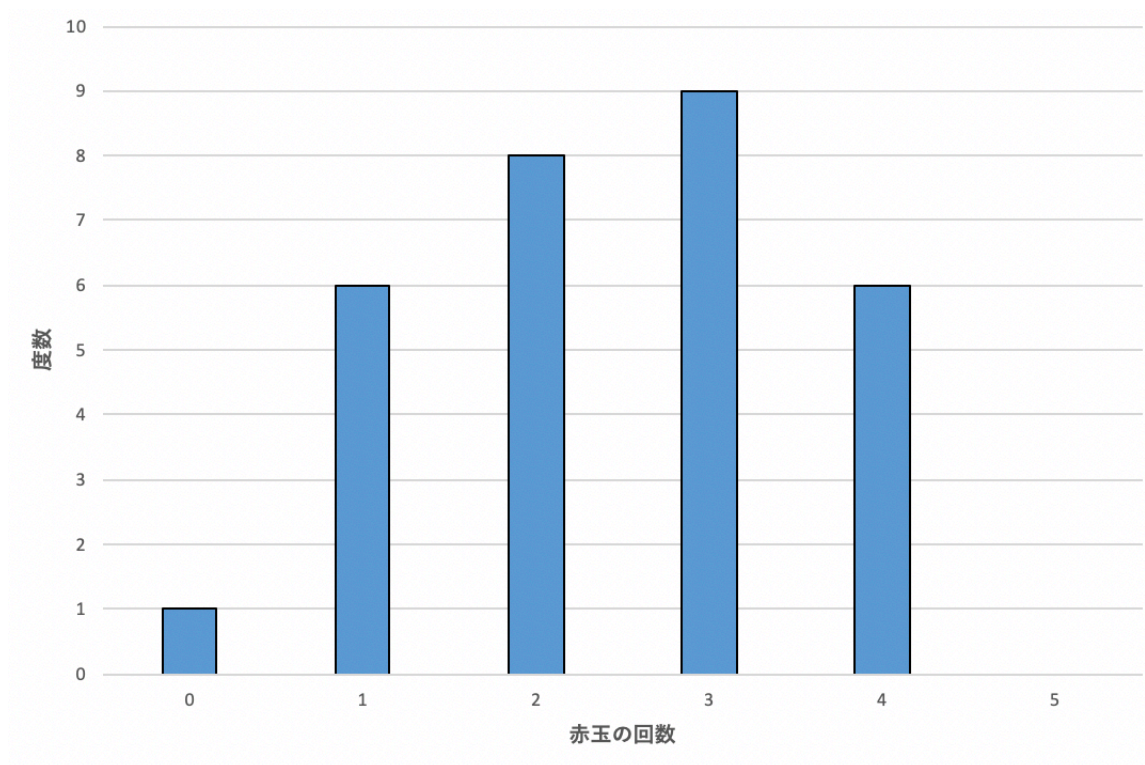


図 12: 東京都の人口の推移

表 3 は、文献 [2] より東京都の 1900 年から 2018 年の人口を示した表である。これをもとにグラフを作成する。時間の推移に対して人口のデータがあるので、折れ線グラフが適切であると考えた。したがって、図 6 に東京都の人口の推移を折れ線グラフで示した。

図 6 を見るとほとんど増加傾向であるが、1945 年前後で急激に低下している。これは、日本が戦争を行っていたので、戦死等の理由から人口が減少したと考える。また、2000 年以降に人口増加が大きいので、経済成長により、都市部に人口が集中していると考える。

5 結論

実験で得られたデータがばらつくことを理解した。また、データを視覚化する方法を学び、データの要約指標のうち代表的なものを学んだ。

6 感想

グラフを読み取ったり考察するのは楽しかったが、エクセルのデータを `tex` の表に変換する作業がとても大変だった。

参考文献

- [1] 東京理科大学工学部情報工学科「情報工学実験 1 2020 年度」(2020/4/6)
- [2] 東京都統計年鑑 平成 30 年 — 2 人口・世帯
<https://www.toukei.metro.tokyo.lg.jp/tnenkan/2018/tn18q3i002.htm>
最終閲覧日:2020/8/3