

## 2016 年度 森林統計学 [課題 2] 基礎統計値の算出

[課題 1]で集めたデータから基礎統計値を算出し、データの分布について考察する。  
※前回の課題では分布状況を視覚的に捉える事が、今回は数値的に評価する事が要点である。

練習問題 (提出不要だが各自課題に取り組む前に解いてみること)

- 1) 教科書 p.29, 例 2. (c)~(h) [資料 “教科書 p.29, 例 2 のデータの集計例” も参照]
- 2) 教科書 p.33, 問題 28. (c)~(i); ただし(i)については四分位数を元データから求める。
- 3) 教科書 p.34, 問題 29. (c)~(i); ただし(i)については四分位数を元データから求める。

※データファイルを含めて過去の資料は下記の鈴木 Web からたどってダウンロード可能

(<http://www.cc.kochi-u.ac.jp/~vsuzuki/kougi.htm#STATISTICS>)

課題 2-1(配点 7 点):

- ① [課題 1]として自分で集めたデータ (2 組以上; データの組数およびデータの個数が多い方が、その苦労を考慮して高く評価する) のそれぞれについて、以下の基礎統計値を算出する (新たに集めたおしなデータを用いてもよい; その場合は[課題 1]と同様の集計・考察もしておくこと)。
  - 平均値  $\bar{x}$  (エクセルを使う場合、合計の算出に SUM 関数を用いてよい。ただし、この課題では平均値は AVERAGE 関数は使わずに求める。) [計算の過程として  $n$  と合計も明記]
  - 標準偏差  $s$  (SUM 関数以外の関数を使わず、教科書 p.21, (9) 式で「分類されていないデータ」として算出する; Web 資料 “教科書 p.29, 例 2 のデータの集計例”の P~U 列の表を参照。) [計算の過程 (計算表を明記) を説明する際に、偏差平方和と分散も明記すること]  
※平均値と標準偏差は、正確な値を得るためには「分類されていないデータ」(個々のデータ, 元データ)を用いて求めなければならない。度数分布表から求める方法 (教科書 p.21, 表 4 の方法) では精度が落ちる (ただし、分類されたデータしかない場合にはこの方法を用いる必要がある)。
    - ・  $\bar{x} \pm s$ , および  $\bar{x} \pm 2s$  の範囲に入る度数とその全体に対する割合を算出 ( $s$  の意味の確認; 余裕があれば  $\bar{x} \pm 3s$  も) [度数分布表からでなく、該当範囲に入るデータの個数から算出]
  - 最大値, 最小値, 範囲 (以上は前回の課題で算出済み), 中央値, 四分位数 (第 1~第 3), 最頻値 (「最頻階級」として最頻階級の階級値を用いる)

- ② 算出した基礎統計値を参考に、データの値の分布状況について考察し、その結果を記述する。

注) 分布形状について比較考察するために、前回作成したヒストグラムも使用して考察を行う (今回のレポートにも再掲する)。なお、ヒストグラムの図に以下の追記をすると統計学的により完全な図となる: まず平均を示す線を書き込み、さらに図の余白に  $n$  (データ数),  $\bar{x}$  (平均値),  $s$  (標準偏差) を記入する。

課題 2-2(配点 3 点): 分布形状の典型的な例として、(a) 教科書第 2 章 6.例 2, (b) 教科書第 2 章 6.例 1, (c) 教科書第 2 章 図 5 (p.16) の 3 つの分布の形状と自分のデータの形状を比較考察する。

注) (a), (b), (c) の順に、釣鐘型 (左右対称で中央付近の度数が多い) から左右非対称でひずみが大きい形状となっている。まずヒストグラムで視覚的に比較し、あわせて①で算出した統計数値との関係を確認する。ひずみの有無・分布の中心はどこか、と、平均値・中央値・最頻階級・四分位数 (最小・最大値も含めて四分位数間の値の差) の関係を(a)~(c)の例と付き合わせて、統計数値と分布形状との関係を理解する。また、自分のデータの分布は(a)~(c)のどの分布形に近いのか (あるいはどれにもあてはまらないか) を、視覚的・数値的根拠も示して考察する。[(c)の  $\bar{x}=9.8, s=11.2$ , 最小値 0, 第 1 四分位数 2.8, 中央値 6.0, 第 3 四分位数 11.0, 最大値 63 とする]

※課題はエクセル等を使用して行なってよい (むしろ推奨する) が、レポートとして形式を整えて必ず A4 用紙の印刷体で提出すること (エクセルの表などを形式を整えずにそのまま印刷しただけでは、課題が正しく行なわれているかどうか判断できないため; ワードにまとめてもよいし、適宜手作業で切り貼りしてもよい)。また基礎統計値の算出については、特に今回は理解度がチェックできるように、計算の過程がわかるように示すこと (計算表を明記する等)。

自習として薦める教科書の問題 (第 2 章 p.31~p.34)

- ・ ヒストグラムについて: 8., 9., 10.
- ・ 平均と標準偏差の計算: 11., 12., 13., 14.
- ・ 平均と標準偏差の理論的背景: 17., 20., 21.
- ・ 中央値・範囲・四分位数: 25., 26., 27.

※第 3 章「確率」は高校数学の範囲なので授業では取り扱わない。不安なものは自習しておくこと。特に「加法定理・乗法定理」は第 4 章の理解に、「順列・組合せ」は第 5 章の理解に必要。