

## 2010年度 森林統計学 [課題3] 2項分布の確率分布

目的: 2項分布がどのようなものであるか(なりたちと性質)を理解する。

基本課題 (7点): まず2項分布に関する教科書の章・節(第5章1.~2.節; p.93~101)を熟読する。

- ① (1) 1回の結果が2種類しかないような独立試行を各自で設定する。[レポートではどのような独立試行か、どういう考えのもとにその試行に決めたか、などを十分説明すること]

例: じゃんけんの勝ち負け、サイコロで1の目が出るか否か、見かけた車のナンバーが県内ナンバーか否か、  
天気が雨か否か、買い物をしたときの金額の端数1円の桁が0か否か、など。

- (2) 独立試行の成功確率  $p$  を、11回以上の試行を繰り返す実験を行なって決定する ( $p =$  成功回数 / 試行回数 [= 11]; 小数点以下3桁に丸める)。11回の試行の結果は表にする。[実験過程を説明]

※ 試行回数は、ある程度信頼性のある平均値を求めるためにはおよそ30回以上のデータが必要であるという経験則によると、本当は30回以上がよい。11回と中途半端な回数にしているのは、分母を素数とすることで、きりのよい計算結果でなくするため(課題なのであえてそのようにしている)。

- ②  $n = 4$  として、成功回数  $x$  の理論分布(2項分布)と経験分布を求める。[計算と実験の過程を説明]

※ 例えば「じゃんけんの勝ち負け」ならば4回のじゃんけんで勝つ回数が  $x$  である(「サイコロの1の目」なら同時にサイコロを4個振って1の目が出た数が  $x$ )。知りたいのは  $x$  で、1回の独立試行の確率  $p$  がわかっているならば  $x$  の発生確率  $P\{x\}$  は2項分布の理論からあらかじめ知ることができる。まずこれを理論分布として求める。そして実際にどのような結果が得られるかの実験を行ないその結果から経験分布を作成し、理論分布(推定値)の検証をする。

- (1) 理論分布: 2項変数  $x = 0, 1, 2, 3, 4$  のそれぞれに対する確率  $P\{x\}$  を求める(教科書 p.97 公式(1); 2項係数の算出は p.293 の付表 II を用いてよい)。平均  $\mu$  と標準偏差  $\sigma$  を、教科書 p.100 ~ 101, 表 4, 表 5 の形式で算出する(教科書 p.101 の(2)式は検算に用いる)。

※ 理論分布は母集団とみなされるので平均と標準偏差には母平均と母標準偏差の記号( $\mu, \sigma$ )を用いる

- (2) 経験分布: 30回以上の試行(実験)を行い、その結果から度数分布と相対度数分布を求める。この相対度数分布が、実験結果による経験分布  $P'\{x\}$  となる。経験分布についても平均  $\bar{x}$  と標準偏差  $s$  を算出する。

※ 「じゃんけんの勝ち負け」ならば4回のじゃんけんで勝つ回数が  $x$  で、これが一組の試行となる。「4回のじゃんけんで買った回数」を30個以上得るために、「4回のじゃんけん」を30回以上行なう。

- ③ 理論分布と経験分布の比較

- (1) 理論分布と経験分布を図にする。

- (2) 理論分布と経験分布のそれぞれについて、確率変数が平均値  $\pm$  標準偏差、平均値  $\pm$  (標準偏差  $\times 2$ ) の範囲に入る確率を求める。

- (3) (1), (2) を参考にして、理論分布と経験分布の相違について考察する(相違はあるか・どの程度か・考えられる原因は何か、等)。

- ④ 自主課題(3点):  $n$  がより大きい場合(例えば  $n = 10$ )の②、③を行い、 $n = 4$  の場合と比較考察する。

※  $n$  が大きくて経験分布を得る実験が大変な場合は、複数の  $n$  (例えば  $n = 10, 20, \dots$  等) の理論分布を作成して、理論分布についてのみ  $n = 4$  の場合と比較考察してもよい。

注) 今回の課題は提出期限を翌々週(6月3日)とする。添付ファイルとしてメールで提出してもよい(提出先: [ysuzuki@kochi-u.ac.jp](mailto:ysuzuki@kochi-u.ac.jp)) が、レポート形式として A4 用紙で印刷可能な体裁のファイルにしておくこと(エクセルファイルそのままでもワード文書に必要な図表を貼り付けるなどの編集を施す)。

自習を薦める教科書の問題 (第5章1~2節 p.117~p.118)

- ・ 2項分布について: 6., 9.
- ・ 独立試行について: 10., 11.
- ・ 2項分布の平均値と標準偏差: 14., 15.