

# 第1講 推測統計の基礎

田中重人 (東北大学文学部准教授)

[テーマ] 推測統計の基礎

## 1 復習

- 記述統計と推測統計 (教科書 pp. 3-5)
- 母集団と標本
- 無作為抽出
- 区間推定と統計的検定の考えかた

## 2 標本比率 $m$ はわかっているが母比率 $M$ が不明の場合の区間推定

つぎのような情報 (= 標本統計量) から、母集団における統計量 (= 母比率) を推測する → 母比率はたぶん ○○ から ×× の範囲にある (区間推定)

袋のなかに色つきの玉がたくさん入っている。ここから8個取り出したところ、すべて赤であった。

この例題では、 $m=1$  であることがわかっているが、 $M$  が不明である ( $n=8$ )。このとき、95%信頼区間を求めるには、 $M$  を適当に仮定し、その仮定の下で  $m=1$  になる確率を計算することを繰り返す：

- もし  $M = 0.9$  なら……
- もし  $M = 0.8$  なら……
- もし  $M =$     なら……

このようにして、 $m=1$  になる確率が **2.5%以上** である  $M$  の範囲を求める。

**課題 1:** 解答を水曜正午までに ISTU に提出。プロセスがわかるように書くこと。

累乗 ( $0.9$  の8乗など) を求めることが必要になる。Windowsの「電卓」ではメニューから [表示] → [関数電卓] に切り替えるとよい。Excelでは $\wedge$ という演算子が使える (掛け算を8回繰り返してもよい)。

### 3 もっと複雑な例

全世界から 400 人を無作為抽出して麺類の好みを訊いたところ、「うどんが好き」と答えた人が 240 人であった。このとき、母集団 (全世界の人々) におけるうどん好きの比率の 95%信頼区間を求めよ (欠損値はないものとする)。

原理的には上記とおなじやりかたで計算できるが、計算量が膨大になるので実際的でない。このような問いに答えるためには、「二項分布」(binomial distribution) の知識を利用する。

### 4 二項分布の簡単な例題

硬貨を 4 回投げて、そのうち表が出る回数  $x$  を数える。

表=1, 裏=0 であらわすと

0 0 0 0 (x=0)  
0 0 0 1 (x=1)  
0 0 1 0 (x=1)  
0 0 1 1 (x=2)  
.....  
1 1 1 1 (x=4)

どれも等しい確率 (1/16) で起こるとすると、つぎのそれぞれの場合の確率が求められる：

表が 1 回も出ない (x=0) 確率：  
表が 1 回出る (x=1) 確率：  
表が 2 回出る (x=2) 確率：  
表が 3 回出る (x=3) 確率：  
表が 4 回出る (x=4) 確率：

**課題 2:** 解答を水曜正午までに ISTU に提出。プロセスがわかるように書くこと。

### 参考資料

- Wikipedia の「二項分布」の項 <<http://ja.wikipedia.org/wiki/二項分布>>
- 高校までの数学の教科書で、順列・組合せと確率・統計をあつかった部分