

システムプログラミング実験

確率プログラミング

2023 年度版
担当：中嶋一貴

実験課題（第 4 回）

第 3 回までの課題を通して、大数の法則は経験的な平均が数学的な平均（期待値）に収束することを確認した。本実験ではさらに中心極限定理（Central limit theorem）について考える。中心極限定理を非常に簡単にいうと、ある n 個の確率変数 X の平均 \bar{X} は n が十分に大きいとき、近似的に平均 μ 、分散 σ^2/n の正規分布に近づいていくというものである。（ここで使われる平均 μ と分散 σ^2 は確率変数 X の母集団の分布の平均と分散を表している。）また、 X の母集団が正規分布に従わない場合においても、十分に大きな n に対して母集団の平均 μ 、分散 σ^2/n の正規分布に従うというものである。

課題 4-1：中心極限定理の確認（母集団が正規分布の場合）

中心極限定理を確認するシミュレーションを作成し、実行せよ。以下の手順に従って取り組むこと。

1. 平均 0, 分散 1 の正規分布に従う乱数 X_i ($i = 1, \dots, n$) を独立に n 個生成し, n 個の平均 $\bar{X}(n) = (\sum_{i=1}^n X_i)/n$ を求める関数 `normal_rand` を作成する.
2. n 個の X の平均 \bar{X}_n を求める試行を m 回行い, m 個の $\bar{X}(n)$ を求める.

なお, 平均 0, 分散 1 の正規分布に従う乱数生成には, c++ の乱数ライブラリ '`std::random`' 内の関数 '`std::random_device`', '`std::mt19937`' および '`std::normal_distribution`' を使用してよい.

レポートには, 以下を記載せよ.

1. $n = 10,000$ とする. $m = 10, m = 100, m = 1,000, m = 10,000$ の 4 通りそれぞれについて, m 個の $\bar{X}(n)$ の平均値と分散. それぞれ有効数字 3 桁で報告すること. また, 理論値との比較を含めること.
2. $n = 10,000, m = 10,000$ としたときの, m 個の $\bar{X}(n)$ のヒストグラム. ヒストグラムの幅は自身で適切なものを定めること. 中心極限定理の結果に基づき, 理論的な分布も一緒にプロットすること.
3. 実験結果の考察.

課題 4-2：サイコロの平均の分布（母集団が一様分布の場合）

第 1 回目のサイコロのシミュレーションを用いて中心極限定理を確認せよ。以下の手順に従って取り組むこと。

1. サイコロを n 回独立に振り、 n 個の出た目の平均 \bar{Y}_n を求める。
2. n 個の出た目の平均 \bar{Y}_n を求める試行を m 回行い、 m 個の \bar{Y}_n を求める。

レポートには、以下を記載せよ。

1. $n = 10,000$ とする。 $m = 10, m = 100, m = 1,000, m = 10,000$ の 4 通りそれぞれについて、 m 個の \bar{Y}_n の平均値と分散。それぞれ有効数字 3 桁で報告すること。また、理論値との比較を含めること。
2. $n = 10,000, m = 10,000$ としたときの、 m 個の \bar{Y}_n のヒストグラム。ヒストグラムの幅は自身で適切なものを定めること。中心極限定理の結果に基づき、理論的な分布も一緒にプロットすること。なお、サイコロの出目の母集団（1 から 6 の整数の一様分布）の平均 μ は $7/2$ 、分散 σ^2 は $35/12$ である。
3. 実験結果の考察。

追加課題 4-A：母集団がベルヌーイ分布の場合

課題 4-1, 4-2 において中心極限定理が母集団に関係なく n 個の平均の分布が母集団の平均と分散/ n に従う正規分布に従うということを確認した。追加課題ではさらに母集団がベルヌーイ分布（二項分布）であることを想定して中心極限定理を確認する。ここでは、ベルヌーイ分布として第 1 回目の課題 1-2 で作成したコイン投げを $p = 0.3$ としてシミュレーションを実行せよ。以下の手順に従って取り組むこと。

1. コイン投げを n 回独立に行い、 n 回のうち、表が出た割合 \bar{P}_n を求める。
2. \bar{P}_n を求める試行を m 回行い、 m 個の \bar{P}_n を求める。

レポートには、以下を記載せよ。

1. $n = 10,000$ とする。 $m = 10, m = 100, m = 1,000, m = 10,000$ の 4 通りそれぞれについて、 m 個の \bar{P}_n の平均値と分散。それぞれ有効数字 3 桁で報告すること。また、理論値との比較を含めること。
2. $n = 10,000, m = 10,000$ としたときの、 m 個の \bar{P}_n のヒストグラム。ヒストグラムの幅は自身で適切なものを定めること。中心極限定理の結果に基づき、理論的な分布も一緒にプロットすること。なお、ベルヌーイ分布の平均と分散については各自で調べること。
3. 実験結果の考察。

実験の感想

レポートの最後に確率シミュレーション実験 4 回を通して感想があれば記載せよ.

レポート提出方法

以下を kibaco の課題ページから提出すること。

- 課題の内容を記載したレポート。
 - － 課題 4-1, 4-2 の内容は必須である。追加課題 4-A の内容は必須ではないが、意欲があれば取り組むと望ましく、加点対象である。
 - － ファイル名は“syspro_pp_xxx_yyy.pdf”とする。xxx には学修番号, yyy には氏名を入力する。例えば学修番号が 22012345 で氏名が Kazuki Nakajima の場合,
“syspro_pp_21012345_kazuki_nakajima.pdf”とする。漢字やひらがなをファイル名に含めないこと。
- 各課題を解くために使用したソースコード。
 - － ソースコードが複数に分かれている場合は、zip ファイルにまとめて提出すること。
 - － 各ソースコードが、誰のもので、どの課題に対応しているかわかるようにファイル名を設定すること。例えば、氏名が Kazuki Nakajima で課題 4-1 に対する C++ コードのファイル名は“kadai_4_1_kazuki_nakajima.cpp”とする。

注意事項

- **本課題の提出締め切りは 1 週間後 (2023/11/22) の 12:00 である。**
- レポート作成時には、テクニカルライティングの資料を確認し、全体をよく検証してから提出すること。
- 提出期限を過ぎると、kibaco から課題を提出できない。提出期限後に課題を提出する場合、下記に連絡をすること。連絡先：nakajima [at] tmu.ac.jp ([at] を@に変える)
- 提出ファイルを間違えていないか、提出前に慎重に確認すること。間違ったファイルが提出されても、教員から学生に逐一連絡することはない。
- 本授業の成績が確定するまで、tmu メールをよく確認すること。成績に関わる重大な事項で教員から連絡する場合がある。