

具体的事象を取り入れた確率・統計の教材構築について 2

On a construction of teaching materials in probability and statistics introduced from concrete events 2

中 村 宗 敬*

NAKAMURA Munetaka

要約：数学的考察の対象となりうる具体的事象として、筆者は以前

- ・ワールドカップやリーグ戦でのサッカーの試合の得点分布
- ・クラス内での誕生日一致ペア出現数を取り上げ、確率論におけるさいころ投げとの関連を探る教材の構築を試みた。それらを実行した主に高等学校生徒対象の出前講義の結果から、大学専門課程以外の学生および高等学校生徒には、サッカーに話題を絞った方が聴講者には理解しやすいとの結論に達した。本小論では、これを修正して非専門的聴講者に対する短時間の講義においても、上記2点の話題を提供することの可能性と残された課題について探求する。

キーワード：さいころ投げ，二項分布，ワールドカップの得点分布，誕生日一致ペアの分布

1 はじめに

[1]において、確率に関連する題材を用いた高校生対象の60分程度の講義の実践報告と、それらを終えてみての見てきた課題を述べた。具体的には、数学的考察の対象となりうる現実事象として、

- ・ワールドカップやリーグ戦でのサッカーの試合の得点分布
- ・クラス内での誕生日一致ペア出現数

を取り上げ、確率論におけるさいころ投げとの関連を探るという内容の話題提供を行った。そこで一つの結論として、非専門的聴講者に対する短時間講義あるいは通常単位構成の15回の一連の講義においても、両方の話題を短い講義時間内に並行的に扱うのは、サッカーに話題を絞った方が聴講者には理解しやすいとの結論に達した（以下では、この結論を「サッカー重点論」と呼び、引用する）。

これは筆者が実践した2つの高等学校における出前講義を分析して得たものである。本小論では、この点の克服の可能性を探る。サッカーと誕生日という、身近に見聞する一見無関係のように見える二つの現実事象が、典型的な確率論の道具であるさいころ投げで結びつけることができることは、多くの人にとって非常に興味をひかれることである。それををなんとか伝えたいという意図は持ち続けており、中途半端な状態を改善する工夫を試みた。その改善点を施した講義・授業も新たに2回行ったので、その報告と振り返り、今後の展望を述べることにする。

まず次の第2節で、の内容を簡単に振り返る。サッカーの得点（数）分布と誕生日一致ペア数の分布が、さいころの1の目の出た回数分布に似ているということから、この三つの類似を探り、サッカーと誕生日が、確率論でいうところの二項分布で特徴づけられる（あるいは近似できるといった方が正確ではある）ということが主内容であった。この内容を盛り込んだ二つの高校での出前講義の実践から、上述したようなサッカーのみという結論に至った経緯をあらためて述べる。

* 科学文化教育講座

次に第3節では、「サッカー重点論」の克服する試みについて述べる。これは、筆者が行っている一般教養科目「確率的見方」を通じて考案し、何回かの講義を通じて行ったものである。それを簡潔に縮約し、一回の講義にも使えるようにした。端的にいうと、サッカー、誕生日一致、さいころ投げの性質を対比できる形で表にまとめあげるという工夫をした。工夫としては、なんということもない些少のものではあるが、比較が容易になるという点では、三事象の共通性の理解を促すことができる。さらに数学から現実への振り戻し段階において、シミュレーション（数値実験）の役割を見直し、コンピューターを使った疑似現実への振り戻しを行うことにした。

次に第4節では、まとめとして統計学習・探求の実践サイクルPPDAC（Problem, Plan, Data, Analysis, Conclusion）の中に第3節の改善点を位置づける。これらを踏まえて、今後の展望を論じる。

2 これまでの振り返り

[1]での実践報告を簡単に振り返る。既に何回か述べている通り、さいころ投げ（8回中）の1の目の回数、誕生日一致ペア数、サッカーの得点数（統一的でないが、取り上げる順序は講義ごとに異なる）の分布に関して、A高では次のような三つのグラフを提示して類似性を指摘した。時系列的にA高の後になるB高では誕生日一致ペア数を削除し、他の二事項の提供にとどめた。

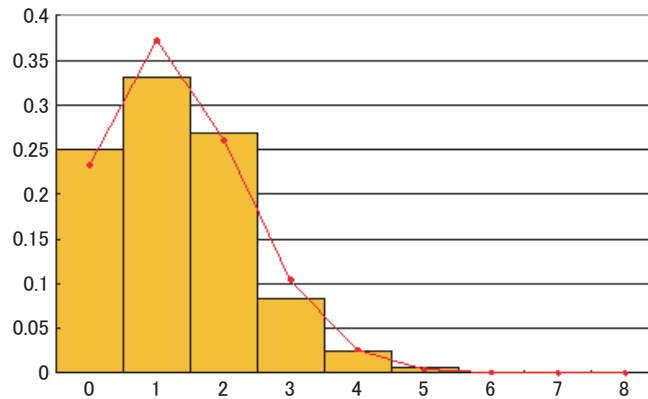


図1 さいころ8回投げシミュレーションによる1の目の回数の相対度数分布（ヒストグラム）と、二項分布 $B(8;1=6)$ （折れ線）

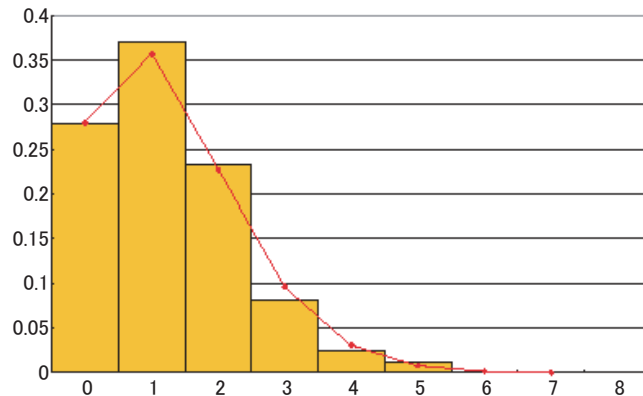


図2 31人集団中の誕生日一致ペア数の相対度数分布（ヒストグラム）と、二項分布 $B(465;465=365)$ （折れ線）。465は $31C2=465$ による

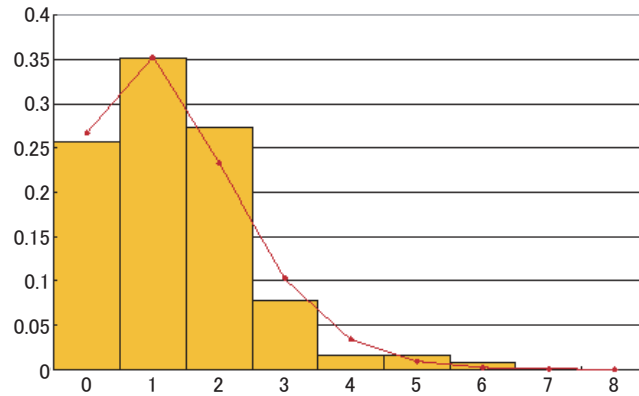


図3 FIFA ワールドカップ 2018 の得点の相対度数分布（ヒストグラム）と、二項分布 $B(90; 1:320=90)$ （折れ線）90 は 1 チーム 1 試合あたりの攻撃機回数で、1.320 は平均得点。よって $1.320/90$ は 1 攻撃機会あたりの得点確率

（[1] 中のグラフを適宜改編して引用した）

A 高、B 高での内容の差異は以下による ([1])。

さて、前節で誕生日一致とサッカー得点分布という 2 つの現実問題を提示したのであるが、高校段階でより適しているのは後者の方であると判断する。クラスの人を誕生日を書き出して同じ日を見つけるという作業を通じて問題を発見するというのは魅力的ではあるが、以下の困難点が見出される：

1. 誕生日の一致が起きるかどうか、また一致ペアが何組生じるかはその場の一回限りのことで、他の集団の状況を想定しにくい。
2. 各ペアでの誕生日一致出現が「ほぼ独立」ということが理解しにくい。一方で、サッカーの方は上の問題点に即して述べると、
 1. 多数のサッカーの試合結果のデータを得ることができる。またテレビ等で実際にいくつもの試合を見ることができる。
 2. 攻撃が終わったら再度 1 からやり直す。すなわち、各攻撃機回はほぼ独立である。

ということになる。加えていうと、誕生日一致があったからといってその意味付けは難しいが、サッカーの結果自体は多くの若い学習者にとっては興味があり、ひいきのチームがあればより関心を持てる。実は前節で A 高の内容を B 高で変更したのは、上述の反省による。

上に掲げた通り、グラフとしては現実に相当するものとして

- ① 誕生日：誕生日生成シミュレーションから得た誕生日一致ペア数分布（誕生日分布は、どの日も等確率であるという標準的なモデル）
- ② サッカー：FIFA ワールドカップ 2018 で開催された 64 試合の得点分布

を提示したのであるが、①は非現実データ、②は現実データという差異があり、非対称な可視化になっている。しかし、1000 集団の誕生日調査は（手作業としては）現実には不可能である。そこで非現実ではありながら、疑似的に現実と見なすことができるという意味でグラフを出したのであるが、今から振り返ると、この点のことで徹底的にできていなかった。現実に起こるものとして、1 回限りの講義受講者の誕生日一致調査に目が向く結果になった。その結果、上記引用の前半の誕生日について言及した 1 のような結論につながった。2 は困難点としては解消していない（数学的に

も非専門者には難しい) が、これもシミュレーション結果を目で見て確かめて、ある程度納得させることが可能であると考えている。

3 サッカー重点論を克服する試み

前節で至った結論、「高校生程度の聴講者にはサッカーのみの方がよい」をなんとか克服したいと考えていた。

一つの方法として、前節で既に半ば述べられているが、シミュレーション結果を現実データと見なすことの強調である。本当の現実である聴講生の誕生日調査は興味喚起のため必要不可欠で、誕生日を扱う際に削除することはできない。そこで、これとの違いと類似性を明確にする必要がある。もう一つは、三つの事象の特徴の共通性を対応する事項ごとに対比させたの表を作成することにより明確にすることである。誕生日調査Aとサッカー(2018J1) B, (8回) さいころ投げCを対比させると次のようになるだろう。

表 A, B, Cの特徴比較

比較項目	A (誕生日)	B (サッカー)	C (さいころ)
①小試行	1 ペアの誕生日の比較	1 回の攻撃機会	さいころ 1 回投げ
②繰り返し回数 N	$\frac{32 \times 31}{2} = 496$	120	8
③小試行間の影響	たぶんなし	ほぼなし	なし
④小試行成功確率 p	$\frac{1}{365}$	$\frac{813}{612 \times 120} = 0.01107$	$\frac{1}{6}$
⑤平均成功回数 Np	$496 \times \frac{1}{365} = 1.359$	$120 \times 0.01107 = 1.328$	$8 \times \frac{1}{6} = 1.333$

注意

- [1] ではサッカー 1 試合 1 チームの攻撃機会数を仮想的に 90 としていたが、実際は平均的に 120 くらいのものである。
- B④の式は、

$$\frac{\text{リーグ全期間の総得点数}}{\text{延べチーム数} \times 1 \text{ 試合 } 1 \text{ チームの攻撃機会数}}$$

を意味している。2018年度J1 (18チーム) での総得点はここにあるように 813, 延べチーム数は 2 試合数 = $2 \times 2 \times {}_{18}C_2 = 612$ である。

⑤の平均がほぼ同じという事項は、グラフが同じような外観になるための必要条件であるためにここに含めておいた。③小試行間の影響がややわかりづらいが、これは措くとして各項目の類似性がこれでかなり容易に把握できるようになったのではないだろうか。

これは、筆者が行っている講義中で作成したものであるが、さいころ投げ以外は空欄にしておき授業出席者とともに、質疑応答を繰り返しながらそれらを埋めていった。データは取れなかったが、効果的な反応が出席者に見られたと考えている。

4 統計学習・探求サイクルPPDACからの展望

本節では、[1]で提示された統計学習サイクルPPDACに関連させて、誕生日・サッカーの問題を述べることにする。

まず、PPDACとは、

Problem (問題発見) → Plan (解決, データ収集のプラン立て) → Data (データ収集実行) → Analysis (データ分析) → Conclusion (結論・評価) → …

の各項目の頭文字をとった統計学習・探求サイクルである。出発点のProblemは、現実から読み取ったごく素朴な形の疑問・問題把握程度のことだと筆者は考えている。データ収集後のAnalysisの段階になってより明確な数学的定式化が図られる。これを誕生日、サッカーにあてはめてみる。

【サッカー】

- [Problem] サッカーはどれくらい点が入るのだろう。サッカーは点が取れにくいスポーツだということを聞くが、本当だろうか。
- [Plan] 実際のデータを調べてみたい。プロのリーグの得点データならば、ネットから得られるだろう。
- [Data] 実際に調べてみる。取得データを見やすい形である得点の度数分布表、ヒストグラムに成型する。
- [Analysis] さいころ投げの確率分布に似ていることに気づき、第3節の表B列の特徴づけから対応する(変形された)さいころ投げモデル化を考え、それに基づいて確率分布の計算を実行する。
- [Conclusion] 前項で得られたモデル化による理論確率分布と得点の度数分布表、ヒストグラムとを照合して、モデル化の適否を判断する。(本来は、統計的検定を行うところであるが、非専門的學生、生徒にはこれは望めない。判断は可視的に行うにとどめる)。

【誕生日一致】

- [Problem] クラス30人の中に同じ誕生日の人たちがいた。これは1年が365日もあることからすると、珍しいのではないか。
- [Plan] 実際のデータを多くのクラスで調べてみたい。だが、現実的にはそれは難しい。コンピュータを使って、シミュレーションを行えば、現実に近い結果が得られるのではないだろうか。
- [Data] 実際にシミュレーションを行ってみる。取得データを見やすい形である一致ペア数の度数分布表、ヒストグラムに成型する。
- [Analysis] さいころ投げの確率分布に似ていることに気づき、第3節の表A列の特徴づけから対応する(変形された)さいころ投げモデル化を考え、それに基づいて確率分布の計算を実行する。
- [Conclusion] 前項で得られたモデル化による理論確率分布と一致ペア数の度数分布表、ヒストグラムとを照合して、モデル化の適否を判断する。(ここも本来は、統計的検定を行うところである)。

以上のように対応づけされた形で、前節の表を参照しながら講義を進行することを計画している。

筆者は今年度も高校生、大学初年次生に講義をする機会を得たが、まだこの内容の効果がどれくらいあるのかというデータをとることまではできていない。まずは、時間的に余裕がある「確率的見方」の方で検証を行ってみて、短時間の出前講義等でもそれを活かすべく精緻化することが今後の課題になる。

参考文献

- [1] 中村宗敬, 具体的事象を取り入れた確率・統計の教材構築について, 山梨大学・教育実践学研究 24, 2019, pp. 115-122.
- [2] フェラー, W., 確率論とその応用, 紀伊國屋書店, 1960.
- [3] ブロム, G., ホルスト, L., サンデル, D., 確率論へようこそ, シュプリンガー・フェアラー東京, 2012.
- [4] 森口繁一, 応用数学夜話, 筑摩書房 (ちくま学芸文庫), 2011.
- [5] Maher, M. J., Modelling association football scores, *Statistica Neerlandica* 36 (1982), nr.314
- [6] Wild, C. J., Pfannkuch, M., Statistical Thinking in Empirical Enquiry, *International Statistical Review* (1999). 67, 3, 223-265.