高校生のための現代数学講座 東京大学
「確率と統計」 玉原国際セミナーハウス
講義 (1)　齊藤　宣一 2016 年 7 月 16 日

「確率と期待値」

この講義では，この講座で皆さんが確率・統計に関する様々な話題を勉強するための準備として，確率の基礎の解説をしたいと思います．講義で扱う概念は，すべて，高等学校の教科書に登場します（ほとんどのことは「数学 A」に書かれています．一部,「数学 B」の内容を扱います）．

まずは，次の問題から始めましょう．1494 年にイタリアで出版された様々な計算技法に関する本（Paciuolo 著）の中では次のような問題が述べられています（[1] の第 10 章）．

問題 1. P さんと Q さんが先に 6 回勝つと賞金 Y 円がもらえるゲームをしている．しかし，P さんが 5 回，Q さんが 2 回勝ったところで，勝負を中断しなくてはならなくなった．このとき，賞金 Y 円をどう分配したら良いであろうか [1]．

数学の計算問題ではないので，絶対的に正しい答えがある問題ではありません．しかし，なるべく両者が納得できるような方法を考えたいと思います．

1. 勝負はついていないので，どちらも 0 円とする．
2. 勝負はついていないので，引き分けと考え，$\frac{1}{2}$Y 円ずつ分配する．
3. 勝った回数の比で分配する．すなわち，P さんには $\frac{5}{7}$Y 円，Q さんには $\frac{2}{7}$Y 円分配する．

「3 でいいや」と考える人が多いかもしれません．しかし，次の問題ではどうでしょうか？

問題 2. P さんと Q さんが先に 16 回勝つと賞金 Y 円がもらえるゲームをしている．しかし，P さんが 15 回，Q さんが 12 回勝ったところで，勝負を中断しなくてはならなくなった．このとき，賞金 Y 円をどう分配したら良いであろうか．

この場合，上記の 3 のように考えると，P さんには $\frac{15}{27}$Y 円，Q さんには $\frac{12}{27}$Y 円を分配することになります．$\frac{15}{27} = 0.55\cdots$，$\frac{12}{27} = 0.44\cdots$ なので，これでは，両者に与えられる賞金の額はほぼ同じになってしまいます．しかし，賞金を全額 Y 円もらえ

[1] 当初配布したプリントでは X 円としていましたが，記号を変えました．

るためには，P さんは，あと一回勝てば良いだけですが，Q さんは 4 連勝しなければなりません．そうすると，3 のように分配したのでは，不公平な感じがします．

カルダーノ[2]はこの問題に対して，$(1+2+3+4):1=10:1$ で分配すれば良いと主張したそうですが，この計算の根拠はよくわかりません．

その後，この問題はパスカル[3]とフェルマー[4]によって（手紙を通じて）詳しく議論されました．現在では，この往復書簡が数学的な確率研究のはじまりとされています．彼らの考えを説明するために，問題を少しやさしくしておきましょう．

問題 3. P さんと Q さんが先に 4 回勝つ（これを優勝すると言うことにする）と賞金 Y 円がもらえるゲームをしている．しかし，P さんが 2 回，Q さんが 1 回勝ったところで，勝負を中断しなくてはならなくなった．このとき，賞金 Y 円をどう分配したら良いであろうか．

この問題に対して，フェルマーは次のように考えたそうです．P はあと 2 回勝てば優勝 ~~です．一方で，Qが優勝するには3連勝するしかありません~~，Q はあと 3 回勝てば優勝です[5]．したがって，少なくとも（多くとも），あと 4 ゲーム行えば，必ず優勝が決まります．そこで，優勝が決まるか否かに関わらず， 4 ゲームを行うことにすると，考えられる勝敗のパターンは，

$$
\begin{aligned}
&\underline{(P,P,P,P)},\\
&\underline{(P,P,P,Q)},\ \underline{(P,P,Q,P)},\ \underline{(P,Q,P,P)},\ \underline{(Q,P,P,P)},\\
&\underline{(P,Q,Q,P)},\ \underline{(P,Q,P,Q)},\ \underline{(P,P,Q,Q)},\\
&\underline{(Q,P,P,Q)},\ \underline{(Q,P,Q,P)},\ \underline{(Q,Q,P,P)},\\
&(P,Q,Q,Q),\ (Q,P,Q,Q),\ (Q,Q,P,Q),\ (Q,Q,Q,P),\\
&(Q,Q,Q,Q)
\end{aligned}
$$

の 16 通りとなります．ここで，例えば，(P,Q,P,Q) は，第 $1,2,3,4$ ゲームの勝者が，それぞれ，P，Q，P，Q であることを意味しています．下線がある場合が，P が優勝する場合です（P が勝つ場合を赤色で，Q が勝つ場合を青色で示しています）．そして，P と Q には実力差がなく，各ゲームにおいて P が勝つことも，Q が勝つことも，“同様に確からしい” と考えました．したがって，上記の 16 通りの各々の場合がおこ

---

[2]Gerolamo Cardano (1501–1576)．イタリアの医師・数学者．三次方程式の根（解）の公式で有名．

[3]Blaise Pascal (1623–1662)．フランスの哲学者・数学者．「人間は考える葦である」．圧力の単位 Pa はこの人の名前からとられた．

[4]Pierre de Fermat (1607?–1665)．フランスの弁護士・数学者．フェルマーの定理（3 以上の自然数 n に対して，方程式 $x^n+y^n=z^n$ を満たす 0 でない自然数 (x,y,z) の組みは存在しない）で有名．

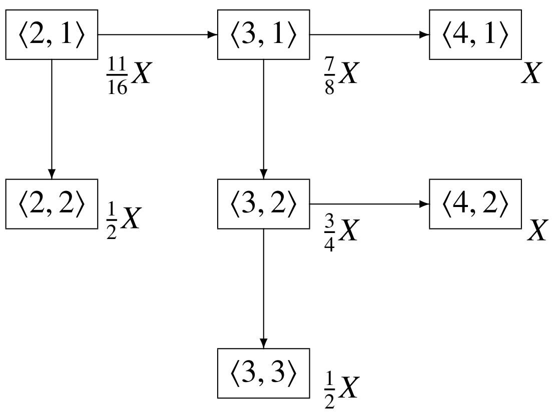
[5]講義当日に配布したプリントに間違いがありましたので，このように修正してください．間違えたことは，消ゴムで消してしまわずに，「自分はこのように間違えた」ことがわかるように残しておくことをすすめます（もちろん間違いが一見してわかるように!）．

ることも“同様に確からしい”ので，P と Q が優勝する場合の数に応じて賞金を分配するべきであると主張しました．したがって，

$$\text{P の賞金は}\frac{11}{16}X, \qquad \text{Q の賞金は}\frac{5}{16}X \tag{$*$}$$

となります．しかし，優勝が決まってからあと 1 ゲーム行うということもあるので，そのようなとき，どちらが勝つのも“同様に確からしい”として良いであろうか？という疑問がでるのは当然でしょう．それはともかくとしても，フェルマーは数学的な考察に基づいて一つの妥当と思われる答えを導きました．

一方で，パスカルは，次のような図を書いて，同じ答え $(*)$ を導きました．この図の中で，例えば，$\langle 2,1\rangle$ は，P が 2 勝，Q が 1 勝の状態を表しています．$\frac{11}{16}X$，$\frac{1}{2}X$ などの数字の意味は，講義の中で説明します．パスカルの考え方は，ある状態から出発した際に，受けとることのできる賞金の平均値を求めていると考えられます．別の言い方をすると（数学の用語を使うと），受けとることのできる賞金の**期待値**を求めています．



ちなみに，このような考え方を適用すると，問題 2 の答えは，P，Q に，それぞれ，$\frac{15}{16}$Y 円，$\frac{1}{16}$Y 円を分配するとなります．

講義では，問題 3 を高等学校で皆さんが学習する確率の知識をつかって考えてみましょう．そのために，確率や期待値の定義を確認して，再度問題 3 に戻り，フェルマーやパスカルの答えを検討します．

参考文献

[1] L. Gårding, Encounter with Mathematics, Springer, 1977.

[2] 高橋幸雄，ゲームの勝敗を確率する，オペレーションズ・リサーチ：経営の科学（公益社団法人日本オペレーションズ・リサーチ学会），41 巻，3 号，1996 年，153–157 ページ.

付録：講義メモ

（ここから先は，講義当日に配布した講義概要にはありませんでした.）

確率の復習のために，次の問題を考えましょう.

問題 4. サイコロを 1 回だけ振ったとき 3 の目が出る確率は？

一つの試行（やってみること）で起こりうる結果の一つを**事象**，事象全体の集合を**全体事象**と言います．事象 A の起こる**確率** $P(A)$ とは，

$$P(A) = \frac{\text{事象 } A \text{ の起こる場合の数}}{\text{起こりうるすべての場合の数}} = \frac{\text{事象 } A \text{ の要素の個数}}{\text{全体事象 } U \text{ の要素の個数}}$$

で定められる数のことです．ただし，根本事象は同様に確からしい，ことを仮定しています．この,「同様に確からしい」とは何だったでしょうか？

サイコロを振った場合，起こりうる結果（事象）は，“1 の目が出る”，“2 の目が出る”，“3 の目が出る”，“4 の目が出る”，“5 の目が出る”，“6 の目が出る” の 6 通りしかありません．そして，“どの目が出やすいということはない”，すなわち，“どの目が出ることも同じ程度に期待” できます．これが，問題 4 では，根本事象は同様に確からしいことに対応しています．したがって，上の定義にしたがって，3 の目の出る確率は

$$\frac{3 \text{ が出る場合の数}}{\text{起こりうるすべての場合の数}} = \frac{1}{6}$$

と計算できます．もう一度確認すると，サイコロを投げる際に，2 の目が出やすいとか 5 の目が出やすいという差異はないことを仮定しています．これは，このような試行を何回も繰り返した際に，1 ~ 6 の目の出る回数はほとんど同じである，という私たちの経験に基づいています（もちろん，本当にこのような試行を経験した人は少ないでしょうが）．このときの，“ほとんど同じ” というのは，全試行回数を n，$1, 2, 3, 4, 5, 6$ の目が出た回数を，それぞれ，$a_n, b_n, c_n, d_n, e_n, f_n$ とした際に，

$$\frac{a_n}{n}, \quad \frac{b_n}{n}, \quad \frac{c_n}{n}, \quad \frac{d_n}{n}, \quad \frac{e_n}{n}, \quad \frac{f_n}{n}$$

の数値（10 進小数）が多くの桁で一致していると言う意味です．$a_n + b_n + c_n + d_n + e_n + f_n = n$ に注意すれば，これら 6 つの分数の値は，n を十分大きくすると，$\frac{1}{6}$ に十分に近い値になるはずです．

以下では（そして，普通は），サイコロを考える時は，このように，“どの目が出ることも同じ程度に期待” できることを仮定します．これは，各目の出る確率が $\frac{1}{6}$ であるようなサイコロを考える，と言っても同じですし，数学的には，より誤解の余地がありません．そして，このような仮定の妥当性は，統計学で検証することにな

ります．したがって，確率と統計は両立されることで，お互いに意味を持つことになります．

続いて，次の問題を考えてみましょう．

問題 5. サイコロを 2 回振って，出た目の合計が偶数のとき「丁」，奇数のとき「半」ということにする．「丁」と「半」のどちらがおこりやすいであろうか？

出る目の合計のパターンは，$2, 3, 4, \ldots, 10, 11, 12$ の 11 通り，このうち，

- 結果が偶数となるのは $2, 4, 6, 8, 10, 12$ の 6 通り
- 結果が奇数となるのは $3, 5, 7, 9, 11$ の 5 通り

となります．したがって，丁となる確率 $\frac{6}{11}$，半となる確率 $\frac{5}{11}$，と考えるのは間違いです．実際，まじめに $6 \times 6 = 36$ 通りを書き出してみると，

$$
\begin{array}{cccccc}
(1,1) & (1,2) & (1,3) & (1,4) & (1,5) & (1,6) \\
(2,1) & (2,2) & (2,3) & (2,4) & (2,5) & (2,6) \\
(3,1) & (3,2) & (3,3) & (3,4) & (3,5) & (3,6) \\
(4,1) & (4,2) & (4,3) & (4,4) & (4,5) & (4,6) \\
(5,1) & (5,2) & (5,3) & (5,4) & (5,5) & (5,6) \\
(6,1) & (6,2) & (6,3) & (6,4) & (6,5) & (6,6)
\end{array}
$$

となり，丁となる確率 $\frac{18}{36} = \frac{1}{2}$，半となる確率 $\frac{18}{36} = \frac{1}{2}$ と計算できます．最初の方法では，出た目の合計が 2 になるのも 4 になるのも同様に扱っていますが，実際には，2 になるのは $(1,1)$ の一通りですが，4 となるのは $(1,3), (2,2), (3,1)$ の 3 通りがあります．これらのことも考慮に入れないと，正しい計算ができないわけです．

次の問題にはどのように答えれば良いでしょうか？

問題 6. サイコロを何回も振った際の出た目の平均値は？

これは，実際に試行を繰り返して，平均を計算するより他ありません．しかし，結果は，試行回数や試行を行う人によって当然変わるでしょう．同じ人が試行した場合でも，今日と明日とでは（試行回数が同じでも）結果は違っていて当然です．そこで，数学的には次のように考えます．先ほどと同様に，全試行回数を n，$1, 2, 3, 4, 5, 6$ の目が出た回数を，それぞれ，$a_n, b_n, c_n, d_n, e_n, f_n$ としましょう．このとき，出た目の平均値は

$$
\begin{aligned}
&\frac{1 \cdot a_n + 2 \cdot b_n + 3 \cdot c_n + 4 \cdot d_n + 5 \cdot e_n + 6 \cdot f_n}{n} \\
&\qquad = \frac{a_n}{n} + 2\frac{b_n}{n} + 3\frac{c_n}{n} + 4\frac{d_n}{n} + 5\frac{e_n}{n} + 6\frac{f_n}{n}
\end{aligned}
$$

となります．先ほど考察したように，各 $\frac{a_n}{n}$ などは，n を十分大きくした際に，$\frac{1}{6}$ に近づくのでした．したがって，平均値も，n を十分大きくした際には，

$$1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} = 3.5 \qquad (\#)$$

と計算でき，ちょうど出る目の真ん中の値になっています（3.5 という目はありませんが）．いちいち，“1 の目が出る”，“2 の目が出る” などと，書いていては面倒なので，それぞれを，$x_1 = 1$，$x_2 = 2$ などと書くことにしましょう（単に，“4 の目が出る” という代わりに “$x_4 = 4$” と書くだけです）．そして，“1 の目が出る確率”，“2 の目が出る確率” などを，p_1，p_2 などと書くことにしましょう．このとき，$p_1 = p_2 = p_3 = p_4 = p_5 = p_6 = \frac{1}{6}$ であることはわかっています．そうすると，(#) は，

$$x_1p_1 + x_2p_2 + x_3p_3 + x_4p_4 + x_5p_5 + x_6p_6 = \sum_{k=1}^{6} x_kp_k$$

と書くことができます．

このように，各事象の起こる確率がわかっているときに，試行の結果の平均値に相当する量を考えることができます．もう少し具体的に表現しましょう．結果が必ず n 個の数値 $x_1, x_2, \ldots, x_n$ のいずれかで表現できるような試行を考えます．このとき，$x_1, x_2, \ldots, x_n$ を一括して一つの記号 X で表現して，例えば，“x_3 が起こる” ことを “$X = x_3$” と表現します．このような X のことを**確率変数**と言います．そして，確率変数の各値 $x_1, x_2, \ldots, x_n$ が起こる確率を，それぞれ，$p_1, p_2, \ldots, p_n$ とします．

確率変数 X	x_1	x_2	x_3	x_4	$\cdots$	x_n
確率	p_1	p_2	p_3	p_4	$\cdots$	p_n

このとき，

$$E(X) = x_1p_2 + x_2p_2 + \cdots + x_np_n = \sum_{k=1}^{n} x_np_n$$

で定められる数 $E(X)$ を X の**期待値**あるいは**平均**と言います[6]．

問題 7. 1 枚 300 円で買える宝くじがある．その宝くじは 1000 万枚（10, 000, 000 枚）につき，

6000 万円：5 本	1500 万円：10 本	1000 万円：10 本
100 万円：60 本	10 万円：595 本	7 万円：90 本
1 万円：2000 本	3000 円：10 万本	400 円：100 万本

[6]数学では，日常的に使う言葉を特定の意味で用いることがよくあります．この平均もそのうちの一つです．確率の話をしている時に，平均と言ったら，それは上記の $E(X)$ のことを意味します．一方で，日常的な意味での平均を使うこともあります．数学では（実は数学でなくても），言葉の意味を文脈に沿って理解することが重要です．

の当たりくじがある．1 枚買ったさいに当たる金額を確率変数と考える時，その期待値を求めよ．

この問題を考える意図は，もちろん，このような宝くじをたくさん買った際に，その平均値（賞金の平均値）がどのように変化するかを調べたい，ということです．どのくじにも，当たりやすさに違いはないと仮定しましょう（そうでなければ，インチキです）．そうすると，例えば，6000 万円が当たる確率は $\frac{5}{1000\cdot 10^4}$ と考えられます．したがって，

$$
\begin{aligned}
E(X) = 6000\cdot 10^4\cdot\frac{5}{1000\cdot 10^4} + 1500\cdot 10^4\cdot\frac{10}{1000\cdot 10^4} + 1000\cdot 10^4\cdot\frac{10}{1000\cdot 10^4} \\
+ 100\cdot 10^4\cdot\frac{60}{1000\cdot 10^4} + 10\cdot 10^4\cdot\frac{595}{1000\cdot 10^4} + 7\cdot 10^4\cdot\frac{90}{1000\cdot 10^4} \\
+ 10000\cdot\frac{2000}{1000\cdot 10^4} + 3000\cdot\frac{10\cdot 10^4}{1000\cdot 10^4} + 400\cdot\frac{100\cdot 10^4}{1000\cdot 10^4} = 139.58
\end{aligned}
$$

と計算できます．すなわち，この宝くじをたくさん買った時の当選金額の平均値は，おおよそ，140 円と考えられますので，1 枚につき 300 円を払ったとすると，元がとれません．そして，何枚買ったところで，この値は変わらないのです．

さて，あらためて，問題 2 を考えてみましょう．まず，$\langle 2,1\rangle$ の状態から，P が優勝する確率を求めてみましょう．例えば，このままゲームを中断せずに続けた際に，(P,P) となれば P の優勝です．あと 2 試合するので 4 通りの結果があり得ます．したがって，(P,P) が実現する確率は $\frac{1}{4}(=\frac{1}{2}\cdot\frac{1}{2})$ となります．また，$(\mathrm{Q},\mathrm{P},\mathrm{Q},\mathrm{P})$ でも P の優勝ですが，これが実現する確率は，$\frac{1}{16}(=\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2})$ となります．このようにして，P あるいは Q が優勝するパターンをすべて書き出して，その確率を計算すると，次のようになります．

P が優勝

番号	勝敗パターン	確率
1	(P, P)	$\frac{1}{4}$
2	(Q, P, P)	$\frac{1}{8}$
3	(P, Q, P)	$\frac{1}{8}$
4	(Q, Q, P, P)	$\frac{1}{16}$
5	(Q, P, Q, P)	$\frac{1}{16}$
6	(P, Q, Q, P)	$\frac{1}{16}$

Q が優勝

番号	勝敗パターン	確率
7	(Q, Q, Q)	$\frac{1}{8}$
8	(P, Q, Q, Q)	$\frac{1}{16}$
9	(Q, P, Q, Q)	$\frac{1}{16}$
10	(Q, Q, P, Q)	$\frac{1}{16}$

したがって，P の優勝確率は $\frac{1}{4}+2\cdot\frac{1}{8}+3\cdot\frac{1}{16}=\frac{11}{16}$，Q の優勝確率は $\frac{5}{16}$ となります．

さて，上の各パターンは一つの試行の結果です．そこで，各番号に対して，

$$x_1 = x_2 = x_3 = x_4 = x_5 = x_6 = \mathrm{Y}, \quad x_7 = x_8 = x_9 = x_{10} = 0$$

と定め，これらを値としてとるような確率変数 X を考えます．これは，P が優勝するか否を問題にしていることになります．それぞれの実現する確率は，上の表から，

$$p_1 = \frac{1}{4}, \quad p_2 = p_3 = p_7 = \frac{1}{8}, \quad p_4 = p_5 = p_6 = p_8 = p_9 = p_{10} = \frac{1}{16}$$

となります．したがって，期待値は，

$$E(X) = \text{Y} \cdot \frac{1}{4} + \text{Y} \cdot \frac{1}{8} + \text{Y} \cdot \frac{1}{8} + \text{Y} \cdot \frac{1}{16} + \text{Y} \cdot \frac{1}{16} + \text{Y} \cdot \frac{1}{16} = \frac{11}{16}\text{Y}$$

となります．

フェルマーの考え方は，優勝確率に応じて賞金を分配するというものであり，パスカルは，この期待値を巧妙な方法で考えたと言えます．彼らの説明は非常に説得力があります．しかし，彼らが手紙を交わしていた時代には，確率という概念も期待値という概念もありませんでした．現在，私たちは，このような概念を使って，比較的機械的に妥当な答えを出すことができます．当然のことですが，このような数学的概念は，はじめからあったわけではなく，先人たちが一つ一つ築き上げてきたものなのです．

— 以上 —