高校生に向けた確率統計の身近な話題

嶋野　和史*

Familiar Topics in Probability and Statistics for High School Students

Kazufumi SHIMANO

要旨　新しい学習指導要領の下で，高等学校数学の在り方が大きな転換期を迎えようとしている．実社会の中に活用されている数学を学ぶ機会が増えることが期待されるが，果たしてこの形で生徒たちが数学をより身近に感じてくれるものかどうかはなかなか厳しいと考えている．高等学校からの依頼を受けた模擬授業に参加する生徒，大学の講義に参加する学生の中に見え隠れする数学への関心がどこにあるのかを考察しながら，今の高校生により身近に感じてもらえる数学の分野として確率統計に注目する．ここでは確率統計の話題を取り上げ，高等学校数学の範囲内でどれだけ身近に関わる数学を知ることができるのかを実感できる工夫について論じていく．

キーワード：確率統計　数学活用　クーポンコレクター問題　ベイズの定理

１．はじめに

高等学校数学に関して，現行の学習指導要領では『数学Ⅰ』『数学Ⅱ』『数学Ⅲ』『数学Ａ』『数学Ｂ』『数学活用』の６科目が対応している．この科目の中で，数学と実生活とのつながりを取り上げているのが『数学活用』となっている．

文部科学省による平成27年度公立高等学校における教育課程の編成・実施状況調査[1)]によると，『数学活用』は全日制の普通科・専門学科では２・３年次に開設している学校はあるものの全体の５～６％前後で，数学の他の科目と比べて極めて低い状況であることが分かる．大学入学試験で試験科目として取り扱われないことが大きく関係していると思われるが，本来の意義でもある数学が如何に身近にあるのかを学ばせるという点で考えれば，全く効果が得られていないと言わざるを得ない．これに関しては，文部科学省の調査からだけではなく，文教大学教育学部に在学している数学専修在籍の学生たちの反応からでも把握していた．著者は2018・2019年度秋学期に開設された『教育課題演習』の科目を担当した．こちらの授業内容は，身近な数学に触れることをテーマとし，高等学校の『数学活用』の教科書[2)]をテキストとした．学生からは「履修したことがない」や「『数学活用』という科目があるとは知らなかった」というものがほとんどであった．

著者が文教大学に着任した2017年より本学の入学センター事務室を通じて，高等学校からの依頼による模擬授業に毎年出向いている．この模擬授業では毎年春学期に開設されている人間科学部・文学部を対象とした科目『総合講座』の著者担当回で取り上げている確率をベースにして授業を行っている．取り上げている話題としては，以下のものである．

①　ジャンケンをしてあいこになる確率
②　クーポンコレクター問題

* しまの　かずふみ　文教大学教育学部学校教育課程数学専修

③　裁判での証言が正しい確率

①については高等学校1年次で履修する『数学A』の単元にある確率の知識だけでも十分対応できるものである．ジャンケンする人数が2人や3人ならば基本問題として取り扱われるものだろうが，これがn人であると少々テクニカルな問題になる．しかし，身近なところに数学があることを知ってもらうための本質がそこにあるとは著者は考えていない．大人数でジャンケンを行うとなかなか勝負が決まらないことは，経験則で誰でも自然に知っていることである．それを数学的な計算で実証できることが，求めるべき本質ではないだろうか．実際に，その部分に立ち入ると，厳密には数列の極限値の知識が必要にはなるが，『数学Ⅲ』の知識があれば十分対応できる．

②は言葉だけでは馴染みがないだろう．しかしながら，子供時代に希望の玩具を取るためにガチャガチャに興じたことはないだろうか．希望の玩具を揃えるためには，お金を払ってガチャガチャに何度か繰り返し挑戦しようとするだろう．それでは，希望の玩具を揃えるために何回挑戦すればよいのだろうというのが，このクーポンコレクター問題の本質である．こちらは残念ながら『数学A』だけの知識では説明できないが，『数学B』の単元である確率分布の話を用いれば対応できる．

③はベイズの定理に関わる問題である．ベイズの定理という形で高等学校の教科書に登場することはないが，発展的な内容として事前の確率と事後の確率との関係で問題が紹介されている．

今回の研究で特に重要視したいのは，②と③が身近にある数学として取り上げられるべき価値が十分高いことと，高等学校数学の多くの科目との関連性が高く，数学の理解を深められる教材といえるということである．

2．クーポンコレクター問題

まず，クーポンコレクター問題は以下のものを一般的には指している．

> n種類のクーポンが等確率，つまり確率$\frac{1}{n}$で取り出されるとする．n種類のクーポンを全て揃えるには，平均何回の取り出しが必要か？

ここでは，クーポンコレクター問題が高等数学のどの単元の知識と結びついて，生徒の学ぶ意識を高める効果があるかどうかを考えながら解いていく． $i=1, 2, 3, \cdots$とし，p_iを既に$i-1$種類のクーポンを獲得した上で新たにi種類目のクーポンを取り出す確率とする．このとき，

$$p_i=\frac{n-(i-1)}{n}$$

である．この計算は，『数学A』でももちろん学ぶが，中学校でも学ぶ基本的な確率の知識からでも求められるだろう．重要なのはここからで，n種類全てのクーポンを揃えるために要した取り出す回数を確率変数として扱うことだ．今，$i-1$種類のクーポンを獲得していたとして，そこから新たにi種類目のクーポンを獲得するまでに要した取り出した回数をX_iとする．$X_i=k$，つまり$k-1$回目までは，今までに獲得したクーポンと同種類のものを取り出し，k回目にして新しい種類のクーポンを取り出したときを考えると，そうなるような確率$P(X_i=k)$は

$$P(X_i=k)=(1-p_i)^{k-1}p_i$$

で計算されることになる．この確率分布は一般的には「幾何分布」として知られているものである．

それでは，すべてのn種類のクーポンを揃えるために要した回数はどうなるかと言えば，

$$X_1+X_2+\cdots+X_n=\sum_{i=1}^{n}X_i$$

となるのである．これをXとおくことにし，確率変数の和であるXの平均（または期待値という）を求めればよいことになる．確率変数の平均については，『数学B』の単元内で学ぶことになる．確率変数の線形性により，

$$E[X]=E[X_1]+E[X_2]+\cdots+E[X_n]=\sum_{i=1}^{n}E[X_i]$$

となる．$E[X_i]$の計算だが，平均の定義から，

$$E[X_i]=\sum_{k=1}^{\infty}k(1-p_i)^{k-1}p_i \tag{2.1}$$

となり，無限に続く和，すなわち，『数学Ⅲ』で学ぶ無限級数の話題と関係することになる．無限和になるのは，クーポンの取り出し方によっては永久に新しい種類のクーポンが現れないこともあり得るからである．この無限級数が収束するかどうかについてだが，以下のような流れで確認することは可能である．

(2.1) の両辺に $1-p_i$ を掛けると，

$$\begin{aligned}(1-p_i)E[X_i]&=\sum_{k=1}^{\infty}k(1-p_i)^{k}p_i\\&=\sum_{k=2}^{\infty}(k-1)(1-p_i)^{k-1}p_i\end{aligned} \tag{2.2}$$

となるので，(2.1)－(2.2) より，

$$p_iE[X_i]=p_i+\sum_{k=2}^{\infty}(1-p_i)^{k-1}p_i$$

となる．上の右辺の第２項は初項 $(1-p_i)p_i$，公比 $1-p_i$ の等比数列の無限和となるので，

$$\frac{(1-p_i)p_i}{1-(1-p_i)}=1-p_i$$

である．以上から，

$$E[X_i]=\frac{1}{p_i} \tag{2.3}$$

が得られる．

(2.3) の結果から，

$$\begin{aligned}E[X]&=\sum_{i=1}^{n}\frac{1}{p_i}=\sum_{i=1}^{n}\frac{n}{n-(i-1)}\\&=n\left(\frac{1}{n}+\frac{1}{n-1}+\cdots+\frac{1}{2}+\frac{1}{1}\right)=n\sum_{i=1}^{n}\frac{1}{i}\end{aligned}$$

と答えが導き出せる．具体的な数値の計算は，小林道正氏の書籍[3] を参照して欲しい．

この問題を生徒の教材として取り上げる利点は，身近な話題を高等学校数学内の科目で十分解き明かせることができることである．数学的なテクニックの難しさは感じることだろうが，高校生が「自分たちの持っている知識」がどのように活用されるのかを実感させることに意味があるのではと考える．

３．等確率ではないクーポンコレクター問題

前章では等確率の場合のクーポンコレクター問題について扱った．数学に関心を持っている生徒なら，

「では，等確率ではない場合はどうなるの？」と自然な疑問が湧くことだろう．数学の学術研究として，等確率ではない場合の問題は，P. Flajolet-D. Gardy-L.Thimonierの1992年の論文[4] において取り上げられている．結果としては次の定理になる．

定理　n 種類のクーポンが取り出される確率がそれぞれp_1，p_2，…，p_nであるとする．n 種類全てのクーポンを揃えるために取り出しに要する平均回数は，

$$\sum_{q=0}^{n-1}(-1)^{n-1-q}\left(\sum_{|J|=q}\frac{1}{1-P_J}\right)$$

である．ただし，J は集合 $\{1,2,\cdots,n\}$ の部分集合とし，

$$\sum_{|J|=q}$$

は要素の個数が q となる J すべてで和をとるものとする．また，

$$P_J=\sum_{j\in J}p_j$$

とする．

定理の証明は，離散数学の知識が必要となり，大学生でも難しく，テクニカルなものになっている．

2020年度の卒業研究において確率論をテーマにしたことから，教員志望の学生でクーポンコレクター問題に興味を持って取り組んでいたものがいた。前章の等確率の場合は有名な結果であり，基本的な数学の知識で証明できることが新鮮に感じ

ていたようだ．その際に，著者が学生に「等確率ではない場合は高校数学で証明することはできないのだろうか？」と質問した．学生自身も色々と調べたようだが，そのような結果はなかなか見つからなかった．そこで著者自身も色々と調べたが，同様に情報が得られなかった．そこで一般の n 種類の場合ではなく，３種類の場合で解決し，上記の定理の結果に辿り着けないか考えることになった．

問題に一緒に取り組んでくれた2020年度ゼミ生の信近氏の卒業論文[5]にも結果が掲載されているので参照して欲しい．ここでは，前章と同様に高等学校数学の単元との関連も述べた上で解答を記述していきたい．

まず，３種類のクーポンに名前をA，B，Cとつけ，取り出される確率がそれぞれ p_1，p_2，p_3 であるとする．上記の定理より，すべてのクーポンを揃えるために要する平均回数は，

$$\sum_{q=0}^{2}(-1)^{2-q}\left(\sum_{|J|=q}\frac{1}{1-P_J}\right)$$

となる．これをΣの記号を使わずに書き下すと，

$$1-\frac{1}{1-p_1}-\frac{1}{1-p_2}-\frac{1}{1-p_3} +\frac{1}{1-p_1-p_2}+\frac{1}{1-p_2-p_3}+\frac{1}{1-p_3-p_1}$$

となる．この章での目標は，高校数学の知識だけで，上の計算結果に辿り着くことである．

前章の等確率の場合では，例えば数種類のクーポンを揃える確率は，どのようなクーポンを取り出したかに依らず計算することができた．等確率ではない場合の重要なポイントは，この部分であるといえる．

３種類のクーポンの揃え方は，A，B，Cの３つの文字列の並べ方であるから，順列の計算により，

$$3!=3\times 2\times 1=6\ (\text{通り})$$

である．その順列のうち，ABC，すなわち，クーポンを取り出すときに，１種類目がA，２種類目がB，３種類目がCという順でクーポンが揃えられた場合に注目しよう．Aが取り出される確率は p_1 である．２種類目のBが取り出されるまでに取り出しに要した回数が m であるならば，Aが取り出された後はAが $m-1$ 回取り出され，m 回目にしてようやくBが取り出されることなる．つまり，２種類目にBが取り出される確率は，

$$p_1{}^m p_2 \qquad (m=1,2,3,\cdots)$$

である．２種類目のBが取り出されてから３種類目のCが取り出されるまでに要した回数が n であるならば，AまたはBが $n-1$ 回取り出され，n 回目にしてCが取り出されることになるから，前に求めた確率を利用することで，クーポンをA，B，Cという順で揃える際に取り出しに要した回数が $1+m+n$ となる確率は，

$$p_1{}^m p_2(p_1+p_2)^{n-1}p_3 \qquad (m,n=1,2,3,\cdots)$$

であることが分かる．ここまでの計算は，前章で触れた幾何分布における確率計算から分かることで，『数学Ａ』の知識で十分理解できる．

それでは，ABC以外の順のクーポンの揃え方ではどうなるであろう？これは，上の計算を繰り返せばよいと気づくことだろう．つまり p の添え字として使っている１，２，３の並び替えを考えたらよい．以上から，３種類のクーポンを全てを揃えるのに取り出しに要した回数が $1+m+n$ である確率は，

$$\sum_{(i,j,k)\in S_3} p_i{}^m p_j p_k (p_i+p_j)^{n-1}$$

で表せる．ここで S_3 は１，２，３の３つの数からなる順列の全体とする．よって，３種類のクーポンを揃えるのに要した平均回数は，

$$\sum_{m+n=2}^{\infty}(1+m+n)\left\{\sum_{(i,j,k)\in S_3} p_i{}^m p_j p_k \left(p_i+p_j\right)^{n-1}\right\} \tag{3.1}$$

となる．この表し方では，高校生には難しく思うかもしれない．しかしながら，(3.1) の和を書き下すことによって，項の並び方に規則性がな

いかどうかを考えさせることは可能ではないだろうか.

(3.1) の和のうち, (i, j, k) を固定した場合を考えてみよう. つまり,

$$\sum_{m+n=2}^{\infty}(1+m+n){p_i}^m p_j p_k\left(p_i+p_j\right)^{n-1} \qquad (3.2)$$

の計算を考えていく. (3.2) について Σ を使わずに書き下すと,

$$\begin{aligned}(3.2) = &3p_ip_jp_k + 4{p_i}^2p_jp_k + 4p_ip_jp_k\left(p_i+p_j\right)\\ &+5{p_i}^3p_jp_k + 5{p_i}^2p_jp_k\left(p_i+p_j\right) + 5p_ip_jp_k\left(p_i+p_j\right)^2\\ &+6{p_i}^4p_jp_k + 6{p_i}^3p_jp_k\left(p_i+p_j\right)\\ &+6{p_i}^2p_jp_k\left(p_i+p_j\right)^2 + 6p_ip_jp_k\left(p_i+p_j\right)^3 + \cdots\end{aligned} \qquad (3.3)$$

となる. (3.3) の式を眺めると, 規則性がある項のかたまりがあることに気づく. 同じ係数となる項たちのうちの 1 項目を順に取り出すと, その和は

$$\sum_{l=1}^{\infty}(l+2){p_i}^l p_j p_k \qquad (3.4)$$

で表せる. これは, 前章の (2.1) の和を求めたときと同様の計算により求めることができる. (3.4) を T とおけば, T を p_i 倍した上で T から引くと,

$$\begin{aligned}(1-p_i)T &= 3p_ip_jp_k + \sum_{l=2}^{\infty}{p_i}^l p_jp_k\\ &= 3p_ip_jp_k + \frac{{p_i}^2p_jp_k}{1-p_i}\end{aligned}$$

となる. ゆえに,

$$T = \frac{3p_ip_jp_k}{1-p_i} + \frac{{p_i}^2p_jp_k}{(1-p_i)^2} \qquad (3.5)$$

である. 次に, 同じ係数となる項たちのうちの 2 項目を順に取り出すと,

$$\sum_{l=1}^{\infty}(l+3){p_i}^l p_jp_k\left(p_i+p_j\right) \qquad (3.6)$$

で表される. (3.6) を p_i 倍したものを (3.6) から引くと,

$$\begin{aligned}&(1-p_i)\times(3.6)\\ &\quad= 4p_ip_jp_k\left(p_i+p_j\right) + \sum_{l=2}^{\infty}{p_i}^l p_jp_k\left(p_i+p_j\right)\\ &\quad= 4p_ip_jp_k\left(p_i+p_j\right) + \frac{{p_i}^2p_jp_k\left(p_i+p_j\right)}{1-p_i}\end{aligned}$$

ゆえに,

$$(3.6) = \frac{4p_ip_jp_k\left(p_i+p_j\right)}{1-p_i} + \frac{{p_i}^2p_jp_k\left(p_i+p_j\right)}{(1-p_i)^2}$$

が得られる. さらに, 同じ係数となる項たちのうちの 3 項目を順に取り出せば,

$$\sum_{l=1}^{\infty}(l+4){p_i}^l p_jp_k\left(p_i+p_j\right)^2 \qquad (3.7)$$

と表され, 上と同様の計算により,

$$(3.7) = \frac{5p_ip_jp_k\left(p_i+p_j\right)^2}{1-p_i} + \frac{{p_i}^2p_jp_k\left(p_i+p_j\right)^2}{(1-p_i)^2}$$

が得られる. この操作を繰り返すことによって, (3.2) は,

$$\sum_{l=1}^{\infty}\frac{(l+2)p_ip_jp_k\left(p_i+p_j\right)^{l-1}}{1-p_i} + \sum_{l=1}^{\infty}\frac{{p_i}^2p_jp_k\left(p_i+p_j\right)^{l-1}}{(1-p_i)^2} \qquad (3.8)$$

で表せる. 注意として, 無限級数の性質として, 値が項の順序の入れ換えや和の取り方に依存しないことを認めるものとする. (厳密な証明は大学生用の微分積分学関連の書籍を探されるとよい.)

(3.8) の第 2 項は, 等比数列の無限和の公式から,

$$\frac{{p_i}^2p_jp_k\big/(1-p_i)^2}{1-\left(p_i+p_j\right)} = \frac{{p_i}^2p_j}{(1-p_i)^2}$$

となる. (3.8) の第 1 項はその形を見れば分かると思うが, 今まで散々繰り返し計算を行っている手法から,

$$\frac{3p_ip_j}{1-p_i} + \frac{p_ip_j\left(p_i+p_j\right)}{(1-p_i)p_k}$$

と求まる. したがって,

$$(3.2) = \frac{3p_ip_j}{1-p_i} + \frac{p_ip_j\left(p_i+p_j\right)}{(1-p_i)p_k} + \frac{{p_i}^2p_j}{(1-p_i)^2} \qquad (3.9)$$

であることが分かった.

（i，j，k）は S_3 の要素として変化するので，(3.1) は，次の6個の和となるはずである．つまり，

・（i，j，k）=（1，2，3）のとき，

$$\frac{3p_1p_2}{1-p_1}+\frac{p_1p_2(p_1+p_2)}{(1-p_1)p_3}+\frac{{p_1}^2p_2}{(1-p_1)^2} \tag{3.10}$$

・（i，j，k）=（1，3，2）のとき，

$$\frac{3p_1p_3}{1-p_1}+\frac{p_1p_3(p_1+p_3)}{(1-p_1)p_2}+\frac{{p_1}^2p_3}{(1-p_1)^2} \tag{3.11}$$

・（i，j，k）=（2，1，3）のとき，

$$\frac{3p_2p_1}{1-p_2}+\frac{p_2p_1(p_2+p_1)}{(1-p_2)p_3}+\frac{{p_2}^2p_1}{(1-p_2)^2} \tag{3.12}$$

・（i，j，k）=（2，3，1）のとき，

$$\frac{3p_2p_3}{1-p_2}+\frac{p_2p_3(p_2+p_3)}{(1-p_2)p_1}+\frac{{p_2}^2p_3}{(1-p_2)^2} \tag{3.13}$$

・（i，j，k）=（3，1，2）のとき，

$$\frac{3p_3p_1}{1-p_3}+\frac{p_3p_1(p_3+p_1)}{(1-p_3)p_2}+\frac{{p_3}^2p_1}{(1-p_3)^2} \tag{3.14}$$

・（i，j，k）=（3，2，1）のとき，

$$\frac{3p_3p_2}{1-p_3}+\frac{p_3p_2(p_3+p_2)}{(1-p_3)p_1}+\frac{{p_3}^2p_2}{(1-p_3)^2} \tag{3.15}$$

となる．

$$\begin{aligned}
&(3.10)+(3.11)\\
&=\frac{3p_1(p_2+p_3)}{1-p_1}+\frac{{p_1}^2(p_2+p_3)}{(1-p_1)^2}\\
&\quad+p_1(p_1+p_2)\cdot\frac{p_2}{(1-p_1)p_3}+p_1(p_1+p_3)\cdot\frac{p_3}{(1-p_1)p_2}\\
&=3p_1+\frac{{p_1}^2}{1-p_1}+p_1(p_1+p_2)\left(\frac{1}{p_3}-\frac{1}{1-p_1}\right)\\
&\qquad+p_1(p_1+p_3)\left(\frac{1}{p_2}-\frac{1}{1-p_1}\right)\\
&=3p_1+\frac{{p_1}^2-p_1(2p_1+p_2+p_3)}{1-p_1}+\frac{p_1(p_1+p_2)}{p_3}\\
&\qquad+\frac{p_1(p_1+p_3)}{p_2}\\
&=3p_1-\frac{p_1}{1-p_1}+\frac{p_1(p_1+p_2)}{p_3}+\frac{p_1(p_1+p_3)}{p_2}\\
&=3p_1+1-\frac{1}{1-p_1}+\frac{p_1(p_1+p_2)}{p_3}+\frac{p_1(p_1+p_3)}{p_2}
\end{aligned} \tag{3.16}$$

と求められ，同様にして，

$$\begin{aligned}
(3.12)+(3.13)=3p_2+1-\frac{1}{1-p_2}+\frac{p_2(p_2+p_3)}{p_1}\\
+\frac{p_2(p_2+p_1)}{p_3}
\end{aligned} \tag{3.17}$$

$$\begin{aligned}
(3.14)+(3.15)=3p_3+1-\frac{1}{1-p_3}+\frac{p_3(p_3+p_1)}{p_2}\\
+\frac{p_3(p_3+p_2)}{p_1}
\end{aligned} \tag{3.18}$$

が求められる．(3.16) - (3.18) により，

$$\begin{aligned}
&\sum_{m+n=2}^{\infty}(1+m+n)\left\{\sum_{(i,j,k)\in S_3}{p_i}^m p_j p_k\left(p_i+p_j\right)^{n-1}\right\}\\
&=6-\frac{1}{1-p_1}-\frac{1}{1-p_2}-\frac{1}{1-p_3}\\
&\qquad+\frac{(p_2+p_3)^2}{p_1}+\frac{(p_1+p_3)^2}{p_2}+\frac{(p_1+p_2)^2}{p_3}\\
&=6-\frac{1}{1-p_1}-\frac{1}{1-p_2}-\frac{1}{1-p_3}\\
&\qquad+\frac{(1-p_1)^2}{p_1}+\frac{(1-p_2)^2}{p_2}+\frac{(1-p_3)^2}{p_3}\\
&=6-\frac{1}{1-p_1}-\frac{1}{1-p_2}-\frac{1}{1-p_3}\\
&\quad+\frac{1-2p_1+{p_1}^2}{p_1}+\frac{1-2p_2+{p_2}^2}{p_2}+\frac{1-2p_3+{p_3}^2}{p_3}\\
&=p_1+p_2+p_3-\frac{1}{1-p_1}-\frac{1}{1-p_2}-\frac{1}{1-p_3}\\
&\qquad+\frac{1}{p_1}+\frac{1}{p_2}+\frac{1}{p_3}\\
&=1-\frac{1}{1-p_1}-\frac{1}{1-p_2}-\frac{1}{1-p_3}\\
&\qquad+\frac{1}{1-p_1-p_2}+\frac{1}{1-p_2-p_3}+\frac{1}{1-p_3-p_1}
\end{aligned}$$

が導き出され，求めたかった結果と一致することができた．

上記の解法は，高校生でも十分理解ができるものであり，何よりも『数学A』の確率，『数学Ⅱ』の分数式の処理，『数学B』の数列，『数学Ⅲ』の無限級数と幅広い視点から考える必要があり，生徒の教材としても非常に適切なものと考える．そして，学術論文の結果に，自分たちの持ち得る知識だけで確認できるということで，数学的な関心を高める効果もあるものと思われる．

４．ベイズの定理

高等学校での模擬授業や大学の教育学部以外の学生対象の講義等の反応を見ると，確率の問題で最も関心があったと思われたのが，ベイズの定理を用いたものだった．しかし，現行の『数学A』の教科書では，ベイズの定理という表現も扱われず，一部の教科書で条件付き確率の応用問題や発展的な内容として紹介されているぐらいである．

ベイズの定理の詳細について説明する前に，ベイズの定理の本質が如何に重要かどうかを確認できる問題を紹介したい．

小林道正氏の書籍[3)]にも登場する以下の問題を考えたい．

> ある町で雨の深夜に起きたタクシーのひき逃げ死亡事故を目撃した人が，「事故を起こしたのは赤色のタクシーでした」と証言した．この証言はどのくらいの確率で信頼できるだろうか？

この情報だけでは数学の問題になり得ないので，以下の数値が分かっているとしよう．

・深夜に走っていたタクシーはほぼ赤色か青色であり，赤色は全体の２割で，青色は全体の８割であった．

・目撃した人の色の識別能力が，赤と青を正しく識別できる確率は80％だった．

この状況の下で，答えを導き出せないものか？

まず，この問題を紹介した直後の生徒や学生の反応は目撃した人の色の識別能力に注目してしまい，目撃した人の証言は，かなりの確率で信用できるものと捉えがちだ．しかし，この判断は否だと言うことに後々気づかされることになる．

証言が信用できる確率とは，目撃した人が赤色のタクシーと証言するという条件で，そのうち本当に赤色のタクシーが事故を起こした割合である．赤色のタクシーと証言する確率は，事故を起こしたタクシーが赤色で，目撃した人がその色を赤と正確に識別した確率と，事故を起こしたタクシーが青色で，目撃した人がその色を赤と誤って識別した確率の和であるから，『数学Ａ』で学習する条件付き確率を用いることで，

$$\frac{2}{10}\times\frac{80}{100}+\frac{8}{10}\times\frac{20}{100}=\frac{320}{1000}=0.32$$

となる．上の計算式から分かるように，正しい証言をしているのは，前半の確率

$$\frac{2}{10}\times\frac{80}{100}=\frac{160}{1000}=0.16$$

である．以上から，証言が信用できる確率は，

$$\frac{0.16}{0.32}=\frac{1}{2}=0.5$$

となる．つまり，目撃した人の証言は信用できるかどうかは判断をつけにくいという結果となる．

上の結果を説明すると，大半の生徒や学生は数学の計算から物事を正しく判断することができるという新しい発見に驚きをもってくれているようだ．

証言から得られたタクシーの色は「事後の事象」，事故を起こしたタクシーの色は「事前の事象」と一般的には捉えられる．この問題の重要なところは，証言の結果がどのような理由でそうなったのかを調べることが可能であることを意味している．これがベイズの定理の利点と言っても良い．

著者の担当する数学専修3年生対象の科目である『確率論』では，ベイズの定理の意味を理解させ，証明も提示している．何よりも大事にしているのは，ベイズの定理の応用が多岐に渡っていることである．度々定期試験や小テストの問題として，次の問題を取り扱っている．

> ある地方では，男性1000人に1人の割合で，ある病気に感染しているという．検査薬によって，感染していれば0.98の確率で陽性反応が出る．ただし，感染していない場合にも，0.01の確率で陽性反応が出るという．さて，いま1人の男性に陽性反応が出たとして，この男性が感染者である確率はどれだけか．

こちらの問題は，著者がある地方の工業高等専門学校教員時代に使用していた教科書[6)]の問題であり，実社会でも活用できる話題として取り上げている問題である．現在のコロナ禍に見舞われている状況を踏まえれば，数学がより身近になっているという実感を湧かすことのできる非常に良い問題と考えている．東京書籍の『数学Ａ』[7)]でも例題としてこれに近い問題が取り上げられているが，高等学校では一部の生徒しか学習できる機会がないと思われる．

こちらの問題は，ベイズの定理を用いた上で解いていく．それではベイズの定理を紹介しよう．

ベイズの定理　事象A_1，A_2が互いに排反とする．Ωを全事象とおくと，

$$A_1 \cup A_2 = \Omega, \quad P(A_1) > 0, P(A_2) > 0$$

とする．このとき，$P(B) > 0$である事象Ｂについて，次が成り立つ．

$$P_B(A_i) = \frac{P(A_i)P_{A_i}(B)}{P(A_1)P_{A_1}(B) + P(A_2)P_{A_2}(B)} \ \ (i = 1, 2)$$

ここで，事象X，Yに対し，$P_Y(X)$は事象Yが起こった下で事象Xの起こる条件付き確率とし，

$$P_Y(X) = \frac{P(X \cap Y)}{P(Y)}$$

と定義するものとする．

上のベイズの定理の記号に合わせて，事象を

A_1：病気に感染している

A_2：病気に感染していない

B ：陽性反応が出る

とする．問題文から，

$$P(A_1) = \frac{1}{1000},\ P(A_2) = \frac{999}{1000}$$
$$P_{A_1}(B) = 0.98,\ P_{A_2}(B) = 0.01$$

である．

$$\begin{aligned} P(B) &= P(A_1)P_{A_1}(B) + P(A_2)P_{A_2}(B) \\ &= \frac{1}{1000} \times 0.98 + \frac{999}{1000} \times 0.01 \end{aligned}$$
$$P(A_1 \cap B) = P(A_1)P_{A_1}(B) = \frac{1}{1000} \times 0.98$$

となるから，求める確率は，

$$\begin{aligned} P_B(A_1) &= \frac{\frac{1}{1000} \times 0.98}{\frac{1}{1000} \times 0.98 + \frac{999}{1000} \times 0.01} \\ &= \frac{0.98}{0.98 + 9.99} = \frac{0.98}{10.97} \\ &= 0.0893345\cdots \ \approx 0.09 \end{aligned}$$

すなわち，陽性反応が出たとして，その男性が感染者である確率は約９％となる．検査薬の精度にもよるが，かなり確率が低いと思われるのではないだろうか．つまり，検査薬によって，いくら陽性反応が出たとしても，それが病気の感染に直結するとは限らないということである．

新型コロナウイルスの感染状況について，ニュースが飛び交っている日々が続くが，陽性率が感染に直結するという判断は，ベイズの定理を理解していれば，場合によっては誤ったものになりかねないという見方ができることから，数学が如何に実生活に役立っているのかを多くの人に知ってもらうきっかけとなり得るものと考えている．

５．まとめ

今回の研究の目標は，高等学校数学の学習内容から，日常の生活と密接したものが数多くあることを，確率統計の分野を通じて再確認することにあった．そして，今回取り上げたクーポンコレクター問題やベイズの定理の応用は，実生活にある問題から数式を立て，計算結果から正しい答えを導くという一番大切な数学を学ぶ意義に触れることができると考える．現行の『数学活用』も，本来はそのような目的で開設されたはずだが，大半の高校生たちに触れられることはなく，正直なところ残念な気持ちになる．

新学習指導要領では，統計教育が拡充され，『数学活用』の内容が単元として盛り込まれるとのことだ．それとは引き換えに，ベクトルの単元の扱いが軽んじられる気配があり，物理の力学分野との繋がりが薄れてしまうのではないかと危惧している．数学と力学との繋がりを知ることは，今回

の研究の本質である数学が身近にあることを実感するまたとない機会になるはずだ．そこが，今後どのような影響をもたらすか注視していきたい．

著者が何より言いたいのは，教科書の内容を教えることは大事だが，それだけでは生徒の理解が深められるわけでもなく，数学を学ぶ意義も伝えることも難しいということだ．数学が如何に大切な学問かを知ることは，本来なら身近なところで幾らでもその機会はあるはずだ．しかしながら，数学科教員側が，生徒たちがその機会に遭遇しやすくする工夫を授業の中でなかなか実践できていないのではと感じている．教員養成に関わる立場として，身近な話題を取り上げた教材開発を進めながら，学生にもその必要性を理解してもらえるように，大学の授業の在り方も大いに検討する時期にありそうだ．

参考文献

１）文部科学省，平成27年度公立高等学校における教育課程の編成・実施状況調査の結果について，https://www.mext.go.jp/a_menu/shotou/new-cs/__icsFiles/afieldfile/2019/02/12/1413569_002_1.pdf

２）岡本和夫他7名，楽しい数学の世界数学活用，実教出版，2012年.

３）小林道正，世の中の真実がわかる「確率」入門，講談社ブルーバックス，2016年.

４）P. Flajolet, D. Gardy and L. Thimonier, Birthday paradox, coupon collectors, caching algorithms and self-organizing search, Discrete Applied Mathematics 39 (1992), 207-229.

５）信近悠太朗，クーポンコレクター問題，令和２年度文教大学教育学部学校教育課程数学専修卒業研究，pp.148-155，2021年.

６）高遠節夫他5名，新確率統計，大日本図書，2013年.

７）俣野博，河野俊丈編，数学Ａ，東京書籍，2015年.