8章 条件付確率

1. 条件付確率

条件付確率は A という事象が起こったもとでの事象 B が起こる確率を $P(B|A)$ と書く。それだけである。単に B が起こる確率ではなく、A が起こっているという条件が付いているということである。これは、以下のように定義できる。

条件付確率

$P(A) \neq 0$ のとき、A と B が両方とも起こる確率を $P(A \cap B)$ とすると、[a]

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

[a] A と B が両方とも起こる確率を $P(A, B)$ と書く場合もある。$A \cap B$ のように集合の記号を使うのは、そもそも厳密な議論で、確率を事象全体の集合の部分集合を $[0, 1]$ の区間内の実数に対応させる写像と定義するからである。

式だけ見ていてもわかりにくいかもしれない。例えば、ランダムにある人 X を連れてくる。一応日本人を連れてきて戸籍を見ることにしよう。そのとき男である確率は、$P(X\text{ は男}) = \frac{1}{2}$ と考えてよいだろう。先ほど戸籍を見ることにしたので $P(X\text{ は女}) = \frac{1}{2}$ でもあるはずだ。ところで、ある日本の大学 (Y 大学としておこう) の男女比が 4 : 1 であったとしよう。さきほど連れてきた X さんはその大学の学生であるとわかったとき、X が男である確率と女である確率どちらが高そうだろうか。もちろん男である確率のほうが高いはずで、5 人に 4 人は男である。条件付き確率で書くと $P(X\text{ は男 }|X\text{ は Y 大学生}) = \frac{4}{5}$ であり、$P(X\text{ は女 }|X\text{ は Y 大学生}) = \frac{1}{5}$ となってほしい。

では、先ほどの式に従って考えてみよう。日本人の人口を N 人 ($N \simeq 1.3 \times 10^8$)、Y 大学の学生数を N_Y 人 ($N_Y = 10000$) とする。[*1] このとき、Y 大学の男子学生数は $\frac{4}{5}N_Y$ 人 $= 8000$ 人だから、

$$\begin{aligned}
P(X\text{ は Y 大学生}) &= \frac{N_Y}{N} \\
&= \frac{10000}{1.3 \times 10^8} \\
P(X\text{ は Y 大学生で、かつ、男}) &= \frac{\frac{4}{5}N_Y}{N} \\
&= \frac{8000}{1.3 \times 10^8}
\end{aligned}$$

よって、

$$\begin{aligned}
P(X\text{ は男 }|X\text{ は Y 大学生}) &= \frac{P(X\text{ は Y 大学生で、かつ、男})}{P(X\text{ は Y 大学生})} \\
&= \frac{4}{5} \times \frac{N_Y}{N} \times \frac{N}{N_Y} \\
&= \frac{4}{5}
\end{aligned}$$

[*1] さきほどは日本の人口や Y 大学の学生数を考えなくてもすんでいたことからも予想できるように、結局 N も N_Y も約分で消えるので細かいことは気にしなくてよい。

2. ベン図での説明

全体集合を U、その部分集合 A と B を考える。1 節の例でいけば U は日本人全体であり、A は日本人男性、B は Y 大学生という感じである。サイコロで言えば、$U = \{1, 2, 3, 4, 5, 6\}$ である。例えば一例だが、以下では A は 4 以上 $A = \{4, 5, 6\}$、B は奇数 $B = \{1, 3, 5\}$ とした。

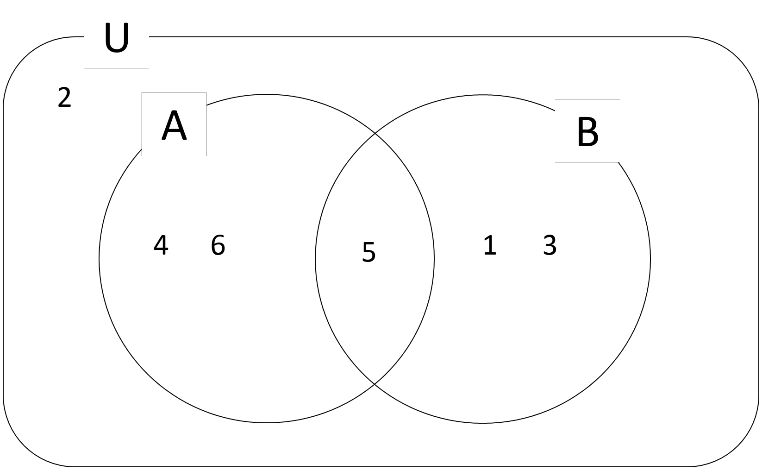

図 1　包含関係を表すベン図

$P(B)$ について考えよう。この確率は全体集合 U のうち、その部分集合 B となる確率を意味している。$U = \{1, 2, 3, 4, 5, 6\}$ で、どの目が出る確率も等しく $\frac{1}{6}$ であるとすると $B = \{1, 3, 5\}$ になる確率は $P(B) = \frac{1}{2}$ である。

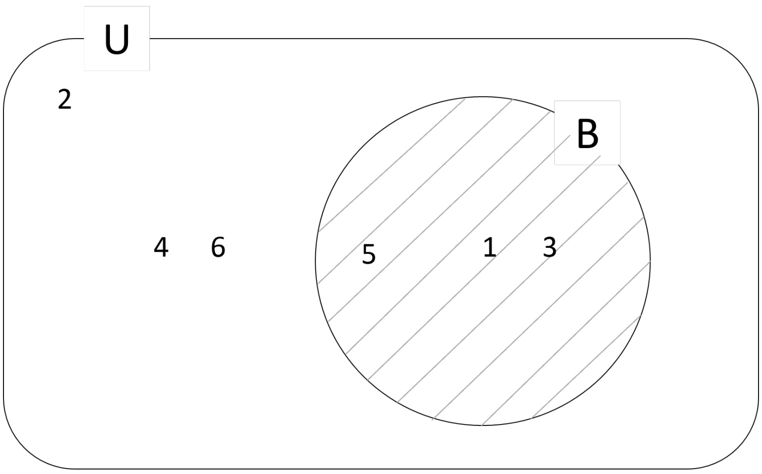

図 2　確率 $P(B)$ は U の中での斜線部 B の確率

では、$P(B|A)$ は同じように考えるとどうなるだろうか。この確率は集合 A のうち、B となる確率、つまり $A \cap B$ となる確率を意味している。条件付き確率 $P(B|A)$ は A であることがすでに分かっている場合なので、A に含まれない場合は無視して良い。そして、A であることは前提として、B である確率ということなのだから A の中での図 3 の斜線部 $A \cap B$ の確率である。$U = \{1, 2, 3, 4, 5, 6\}$ で、どの目が出る確率も等しく $\frac{1}{6}$ で

あるとすると $A = \{4, 5, 6\}$ のうち、$A \cap B = \{5\}$ となる確率は $P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{1}{3}$ である。

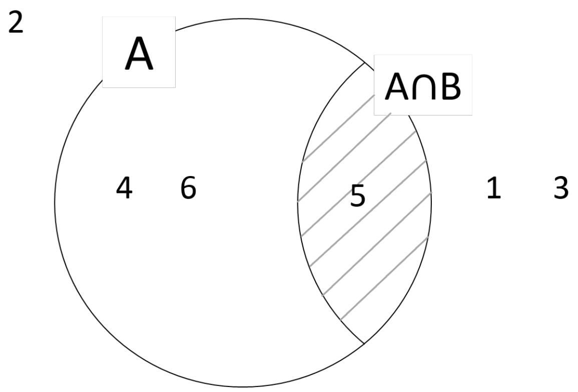

図 3　確率 $P(B|A)$ は A の中での斜線部 $A \cap B$ の確率

つまり、条件付き確率とは A であるという前提が加わることで、もととなる集合が全体集合 U からその部分集合 A へと変わったというだけである。

3. ベイズの定理

条件付き確率を語るうえで決して外すことができない定理がベイズの定理である。とはいっても、条件付き確率の定義からほぼ自明であろう。しかし、これを使うとすっきりと語れる確率があることも間違いない。

ベイズの定理

A と B という事象がある。それぞれの起こる確率を $P(A) \neq 0$, $P(B) \neq$ とし、A が起こったもとでの B が起こる確率を $P(B|A)$、A が起こったもとでの A が起こる確率を $P(A|B)$ とすると以下が成り立つ。

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)}$$

証明しよう。A と B が両方とも起こる確率を $P(A \cap B)$ とすると、条件付き確率の定義から

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$
$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

両辺それぞれ、$P(A)$ 倍、$P(B)$ 倍して、

$$P(B|A)P(A) = P(A \cap B)$$
$$P(A|B)P(B) = P(A \cap B)$$

右辺はともに $P(A \cap B)$ であるから、

$$P(B|A)P(A) = P(A|B)P(B)$$
$$P(B|A) = \frac{P(A|B)P(B)}{P(A)}$$

証明終わり。
とはいえ、理解を深めるには、やはり実際に使ってみるのが一番である。確率の話ではおなじみの袋に入った紅白の球を考えよう。
袋が二つある。外見から見分けはつかないが便宜的に袋 X と袋 Y としよう。袋 X には赤い球が 6 個、白い球が 1 個入っており、袋 Y には赤い球が 2 個、白い球が 3 個入っている。まず、ランダムに一方の袋をとる。そして、中を見ないでランダムに球を一つ取り出した。この球が赤色であったとき、とった袋が X である確率はいくつか。
まず、とった袋が X か Y かどちらの確率が高いと言えそうだろうか。袋 X のほうが入っている赤い球が多いのだから、X のほうが Y よりも確率は高いだろう。

	赤い球	白い球	合計
袋 X	6	1	7
袋 Y	2	3	5
合計	8	4	12

表 1　球の個数

具体的に解いてみる。ベイズの定理から、

$$P(\text{袋 }X|\text{ 球は赤色}) = \frac{P(\text{球は赤色} \mid \text{袋 }X)P(\text{袋 }X)}{P(\text{球は赤色})}$$

右辺の確率をそれぞれ分かれば、求めたい左辺の $P(\text{袋 }X|\text{ 球は赤色})$ もわかる。袋 X だとわかっていれば、球は 7 個入っていて、赤い球が 6 個、白い球が 1 個だから、

$$P(\text{球は赤色} \mid \text{袋 }X) = \frac{6}{7}$$

袋 X を選ぶ確率はランダムに X か Y をとるから、

$$P(\text{袋 }X) = \frac{1}{2}$$

最後が少しわかりにくいかもしれない。$P(\text{球は赤色}) = \frac{8}{12} = \frac{2}{3}$ とするのは誤りである。袋 X と袋 Y に入っている球の数が異なり袋 X のほうが多いので、袋 X と袋 Y を $\frac{1}{2}$ でどちらか選ぶと、袋 X の球と袋 Y の球では取り出される確率が袋 Y のほうが小さくなる。袋が X か Y かわからない状態での球が赤色である確率だから、袋 X を選び赤い球を取り出す確率と袋 Y を選び赤い球を取り出す確率の和である。

$$P(\text{球は赤色}) = P(\text{球は赤色} \mid \text{袋 }X)P(\text{袋 }X) + P(\text{球は赤色} \mid \text{袋 }Y)P(\text{袋 }Y)$$

ちなみに、この式は $P(\text{袋 }X|\text{ 球は赤色}) + P(\text{袋 }Y|\text{ 球は赤色}) = 1$ であるから、ベイズの定理より

$$\begin{aligned}
P(\text{球は赤色}) &= \{P(\text{袋 }X|\text{ 球は赤色}) + P(\text{袋 }Y|\text{ 球は赤色})\}\, P(\text{球は赤色}) \\
&= P(\text{袋 }X|\text{ 球は赤色})P(\text{球は赤色}) + P(\text{袋 }Y|\text{ 球は赤色})P(\text{球は赤色}) \\
&= P(\text{球は赤色} \mid \text{袋 }X)P(\text{袋 }X) + P(\text{球は赤色} \mid \text{袋 }Y)P(\text{袋 }Y)
\end{aligned}$$

としても求められる。

$$P(\text{袋}\,Y)=\frac{1}{2}$$
$$P(\text{球は赤色}\mid\text{袋}\,Y)=\frac{2}{5}$$

なので、

$$\begin{aligned}P(\text{球は赤色})&=P(\text{球は赤色}\mid\text{袋}\,X)P(\text{袋}\,X)+P(\text{球は赤色}\mid\text{袋}\,Y)P(\text{袋}\,Y)\\&=\frac{6}{7}\times\frac{1}{2}+\frac{2}{5}\times\frac{1}{2}\\&=\frac{22}{35}\end{aligned}$$

したがって、求めたい確率 $P(\text{袋}\,X|\,\text{球は赤色})$ は

$$\begin{aligned}P(\text{袋}\,X|\,\text{球は赤色})&=\frac{P(\text{球は赤色}\mid\text{袋}\,X)P(\text{袋}\,X)}{P(\text{球は赤色})}\\&=\frac{5}{6}\times\frac{1}{2}\div\frac{22}{35}\\&=\frac{175}{264}\end{aligned}$$

である。$P(\text{袋}\,Y|\,\text{球は赤色})=1-P(\text{袋}\,X|\,\text{球は赤色})=\frac{89}{264}$ だから、最初の推測通り X のほうが確率が高いことがわかる。[*2]なお、直観的だと思い断りなく用いたが、$P(\text{袋}\,X|\,\text{球は赤色})+P(\text{袋}\,Y|\,\text{球は赤色})=1$ は、ベイズの定理から、

$$\begin{aligned}&P(\text{袋}\,X|\,\text{球は赤色})+P(\text{袋}\,Y|\,\text{球は赤色})\\=&\frac{P(\text{袋}\,X\,\text{かつ、球は赤色})+P(\text{袋}\,Y\,\text{かつ、球は赤色})}{P(\text{球は赤色})}\\=&\frac{P\left((\text{袋}\,X\,\text{かつ、球は赤色})\,\text{または、}(\text{袋}\,Y\,\text{かつ、球は赤色})\right)-P\left((\text{袋}\,X\,\text{かつ、球は赤色})\,\text{かつ、}(\text{袋}\,Y\,\text{かつ、球は赤色})\right)}{P(\text{球は赤色})}\\=&\frac{P\left((\text{袋}\,X\,\text{または、袋}\,Y)\,\text{かつ、球は赤色}\right)-P\left((\text{袋}\,X\,\text{かつ、袋}\,Y)\,\text{かつ、球は赤色}\right)}{P(\text{球は赤色})}\\=&\frac{P(\text{球は赤色})-0}{P(\text{球は赤色})}\\=&1\end{aligned}$$

[*2] 表に出てくる数字をすべて異なる数字にすることでどの数字を使ったのかわかりやすくするつもりだったが、結果として計算が煩雑になったことは申し訳ない。

なお、一般に B の補集合を B^c とし、[*3]全体集合を $U = B \cup B^c$、空集合を $\emptyset$ とすると、

$$
\begin{aligned}
&P(B|A) + P(B^c|A) \\
=&\frac{P(B \cap A) + P(B^c \cap A)}{P(A)} \\
=&\frac{P((B \cap A) \cup (B^c \cap A) - P((B \cap A) \cap (B^c \cap A))}{P(A)} \\
=&\frac{P((B \cup B^c) \cap A) - P((B \cap B^c) \cap A)}{P(A)} \\
=&\frac{P(U \cap A) - P(\emptyset \cap A)}{P(A)} \\
=&\frac{P(A) - P(\emptyset)}{P(A)} \\
=&1
\end{aligned}
$$

$B =$ 袋 X とすれば、補集合 $B^c =$ 袋 Y である。計算は同じことをしているので比較しながら見てもらいたい。

4. 具体的な話題

具体的な例として偽陽性[*4]の話がある。10 万人に 1 人 ($p = \frac{1}{10^6}$) が発症する病気があるとしよう。あなたは、この病気の検診を受けた。すると、結果は陽性[*5]、実際には病気でない人のうち間違って陽性と判断される (偽陽性) のは 100 人に 1 人 ($q = \frac{1}{100}$)。絶望的である。
かというと、そうでもないというのがこの話の肝である。病気の人のうち間違って病気でない (陰性) と判断されてしまうのは、ここでは 1000 人に 1 人 ($r = \frac{1}{1000}$) ということにしよう[*6]。

	陽性	陰性	合計
病気あり	$p(1-r)$	$pr = \frac{1}{10^9}$	$p = \frac{1}{10^6}$
病気なし	$q(1-p) = \frac{1-p}{100}$	$(1-q)(1-p)$	$1-p$
合計	$p(1-r) + q(1-p)$	$pr + (1-q)(1-p)$	1

表 2　実際の病気の有無と検査の結果の確率

病気である確率は p、そのうち間違って陰性になる確率は r だから、病気なのに陰性になる確率は pr、病気でかつ正しく陽性となるのは間違わないと考えれば余事象から $p(1-r)$ といったように計算できる。先ほど絶望的だと思ってしまったのは、病気でないのに陽性と判定される確率 $q = \frac{1}{100}$ のせいである。しかし、条件付き確率を考えてみると $q = P($陽性 $|$ 病気なし$)$ である。今、陽性だとわかっているのだから知りたいのは、陽

*3 補集合といえば $\bar{B}$ かもしれないが、これは集合の閉包として使われることが多いため、こちらの記号にも慣れておくとよい。
*4 文字通り偽の陽性である。
*5 結果が病気であることを示すとき陽性、そうでないとき陰性という。
*6 実際に病気なのに見逃してしまうほうが、病気でない人を誤って病気だと判断するよりもまずいことになるのはわかるだろう。だから、ふつうは $r < q$ である。

性であるとわかったときの病気ありと病気なしの確率である。それは、

$$
\begin{aligned}
P(\text{病気あり} \mid \text{陽性}) &= \frac{P(\text{病気あり、かつ、陽性})}{P(\text{陽性})} \\
&= \frac{p(1-r)}{p(1-r)+q(1-p)} \\
&= \frac{999p}{989p+10} \\
&= \frac{999}{1000989} \\
&\simeq \frac{1}{1000} \\
P(\text{病気なし} \mid \text{陽性}) &= \frac{P(\text{病気なし、かつ、陽性})}{P(\text{陽性})} \\
&= \frac{q(1-p)}{p(1-r)+q(1-p)} \\
&= \frac{10(1-p)}{989p+10} \\
&= \frac{999990}{1000989} \\
&\simeq \frac{999}{1000}
\end{aligned}
$$

なお、余事象なので一方を計算すれば十分である。結果を見ると、実際には $\frac{999}{1000}$ の確率で病気でないことがわかるから、まだ絶望的ではない。

次にモンティホール問題の話をしよう。これも条件き付確率の話題としては有名である。モンティホール問題とは、とある番組で行われていたゲームで、モンティホールはその司会者の名前だそうだ。ルールを整理しよう。

0. 三つの扉があり当たりは一つで残り二つははずれ。(当たりの扉を扉 X としよう)

1. まず、プレイヤーはどれか一つを選ぶ。(見分けはつかないが便宜上これを扉 A とし、残りを B, C とする。)

2. 次に司会者はプレイヤーが選んだ扉 A 以外の二つから一つはずれの扉を選び、それを開けてプレイヤーにはずれであると示す。

3. プレイヤーは選択する扉をかえてもよいし、A のままでもよい。

さて、あたりを引く確率は変更したほうが高いか変更しないほうが高いか、それとも同じか。というのが問題である。実は、2 番目の司会者の操作が肝で、しっかりと定義しないと確率が求まらない可能性があるから、しっかりと手順を確認する。もし、A がはずれ $(X \neq A)$ であれば、B, C のうち一方はあたりであるから、司会者ははずれである方を開けるしかない。B, C のうち B がはずれである確率は $P(X \neq B | X \neq A) = \frac{1}{2}$ である。また、司会者が開ける扉を Y とすると、このとき司会者は B を開けるしかないから、$P(Y = B | X \neq B \text{ かつ、} X \neq A) = 1$ である。

一方、もし A があたり $X = A$ であれば、B, C はどちらもはずれであるから、司会者はどちらを開けてもよい。B, C のどちらもはずれだから、$P(X \neq B | X = A) = 1$ である。このとき確率 $\frac{1}{2}$ でどちらかを開けると

しておくと、$P(Y=B|X\neq B$ かつ$, X=A)=\frac{1}{2}$ である。一般に、事象 X,Y,Z に対し以下が言える。

$$\begin{aligned}P(X|Y\cap Z)P(Y|Z)&=\frac{P(X\cap Y\cap Z)}{P(Y\cap Z)}\times\frac{P(Y\cap Z)}{P(Z)}\\&=\frac{P(X\cap Y\cap Z)}{P(Z)}\\&=P(X\cap Y|Z)\end{aligned}$$

司会者によって開かれれば、はずれだということはわかっているので、$P(X\neq B|Y=B)=1$ だから

$$\begin{aligned}P(Y=B \text{ かつ}, X\neq B)&=P(X\neq B|Y=B)P(Y=B)\\&=P(Y=B)\end{aligned}$$

したがって、

$$\begin{aligned}P(Y=B|X=A)&=P(Y=B|X\neq B \text{ かつ}, X=A)P(X\neq B|X=A)\\&=\frac{1}{2}\\P(Y=B|X\neq A)&=P(Y=B|X\neq B \text{ かつ}, X\neq A)P(X\neq B|X\neq A)\\&=\frac{1}{2}\end{aligned}$$

求めたいのは $P(X=A|Y=B)$ や、$P(X=C|Y=B)$ である。ベイズの定理から、

$$\begin{aligned}P(X=A|Y=B)&=\frac{P(Y=B|X=A)P(X=A)}{P(Y=B)}\\&=\frac{1}{2}\times\frac{1}{3}\div\frac{1}{2}\\&=\frac{1}{3}\end{aligned}$$

$P(X=A)$ は司会者からなんの情報も得る前に選んだ A があたりである確率だから $\frac{1}{3}$ である。$P(Y=B)$ はプレイヤーの最初の選択によらず B が開かれる確率だから、$\frac{1}{2}$ である。納得できなければ、

$$\begin{aligned}P(Y=B)&=P(Y=B|X=A)P(X=A)+P(Y=B|X\neq A)P(X\neq A)\\&=\frac{1}{2}\times\frac{1}{3}+\frac{1}{2}\times\frac{2}{3}\\&=\frac{1}{2}\end{aligned}$$

と考えればよい。余事象から $P(X=C|Y=B)=P(X\neq A|Y=B)=1-P(X=A|Y=B)=\frac{2}{3}$ である。これを余事象によらず求めるのは読者が各自やってみてほしい。結果としては、もともと選んだものをかえるほうが得だというわけである。しかし、これが直観からずれているため当時の数学者の間でももめたらしい。

5. 独立

話題は変わって、確率における独立の意味について書いておく。

事象の独立

事象 A と B が独立であるとは、

$$P(A\cap B)=P(A)P(B)$$

となることである。

サイコロを 1 回振って 1 の目が出る確率は $\frac{1}{6}$ とすると、同じサイコロを 2 回振って 2 回とも 1 が出る確率が

$\frac{1}{36} = \frac{1}{6} \times \frac{1}{6}$ であれば、1 回目に 1 がでる事象と 2 回目に 1 が出る事象は独立である。確率変数が独立だということも、5 章などで使っていたが、それは確率変数がどの値をとる事象も独立ということである。

確率分布のの独立

確率分布 X と Y が独立であるとは、任意の実数 x, y に対し、

$$P(X \leq x \text{ かつ、} Y \leq y) = P(X \leq x)P(Y \leq y)$$

となることである。

ところで、条件付き確率の定義から

$$\begin{aligned} P(A \cap B) &= P(A|B)P(B) \\ &= P(B|A)P(A) \end{aligned}$$

であった。独立を考えると、

$$\begin{aligned} & P(A \cap B) = P(A)P(B) \\ \Leftrightarrow \quad & P(A|B) = P(A) \\ \Leftrightarrow \quad & P(B|A) = P(B) \end{aligned}$$

である。意味を考えると、条件付き確率は前提となる条件によらずもとの確率と等しいことが独立と同値であり、$P(A \cap B)$ が $P(A)P(B)$ と積でかけるという定義よりも、直観的であると思う。