従属である場合

2 つの事象が独立でなければ、それは**従属**と呼ばれる。すなわち、1 つの事象が起こるか起こらないかによって、他の事象の確率が影響を受けることを意味する。端的な例はくじ引きだろう。A、B の 2 人が順番にくじを引くとき、後に引くものは先に引いた者の結果に影響を受ける。もし A が先にくじを引いてはずれたなら、後から引く B は当たる確率が上がる。なぜなら、当たりくじの数は変わらないがくじの総数は 1 本少なくなっているからである。当然、当たりが占める割合は高くなっている。しかし、A が先に当たりを引いたら、後から引く B は当たる確率が下がる。理由は単に当たりが 1 本減ったからではない。くじの総数も 1 本少なくなっているので、ちょっとした計算はした方がよいだろう。

N 本中 a 本が当たりとしよう。この状態で当たる確率は $\dfrac{a}{N}$ である。当たりが 1 本抜かれた後では当たる確率は $\dfrac{a-1}{N-1}$ になる。$\dfrac{a}{N} > \dfrac{a-1}{N-1}$ を示したいのだが、それは

$$\begin{aligned} \frac{a}{N} - \frac{a-1}{N-1} &= \frac{a(N-1)-(a-1)N}{N(N-1)} \\ &= \frac{N-a}{N(N-1)} \end{aligned}$$

であり、この値は間違いなく正である。すなわち、当たる確率は確実に下がるのである。

さて、くじ引きは後に引く者が先に引く者の影響を受けることが分かった。すなわち、従属なのである。では、くじ引きは先に引く者と後に引く者ではどちらが有利なのだろうか。N 本中 a 本が当たるくじについて考えることにしよう。

先に引く A が当たりを引く確率は簡単な話で、それは $\dfrac{a}{N}$ である。B が当たりを引く確率を知りたいのだが、それには 2 通りの状況を考えなくてはならない。A が当たりを引いた後に B も当たりを引く場合と、A がはずれを引いた後に B は当たりを引く場合である。

A が当たりを引いた場合に B が引くくじは、$N-1$ 本中 $a-1$ 本が当たるくじを引くことに等しい。すなわち B が当たる確率は $\dfrac{a-1}{N-1}$ である。A がはずれを引いた場合に B が引くくじは、$N-1$ 本中 a 本が当たるくじを引くことに等しい。すなわち B が当たる確率は $\dfrac{a}{N-1}$ である。よって、

$$\begin{aligned} \frac{a}{N}\cdot\frac{a-1}{N-1} + \frac{N-a}{N}\cdot\frac{a}{N-1} &= \frac{a(a-1)+(N-a)a}{N(N-1)} \\ &= \frac{Na-a}{N(N-1)} \\ &= \frac{a(N-1)}{N(N-1)} = \frac{a}{N} \end{aligned}$$

となって、結局 A が当たりを引く確率に等しいことが分かる。くじ引きは、引く順番に関係なく当たる確率が等しい公平な試行なのである。

*　*　*

この結果から、くじ引きにおいては後から引く者は先に引いた者の影響を受けないことになって、従属と言えないのではないかと思うかもしれない。たしかに結果的に影響を受けないように見えるが、実際は先に引いた影響のために後に引く者の確率が等しくなっているのである。

極端な話として、当たり 1 本だけのくじ、もしくは、はずれ 1 本だけのくじを考えてみよう。このとき

はずれを引かなかった者が当たり

という条件でくじを引けば、先に引く者と後に引く者の確率は明らかに異なる。▮

従属事象の定理

いま、従属事象の確率に関する例でくじ引きを取り上げ、何気なく積の法則を用いたことに注意してもらいたい。一般に事象 X の確率を $P(X)$、事象 Y の確率を $P(Y)$ で表すとすると、X、Y が独立であるなら

$$P(X \cap Y) = P(X)P(Y)$$

成り立つのであった。ところで確率 $P(Y)$ が、X が起こるか起こらないかによって変動するとき、すなわち従属であるとき、X の影響を受けることを $P_X(Y)$ のように明示する場合がある。このとき、Y の確率を

X のもとでの Y の条件付き確率

と呼ぶことにする。

たとえば、袋の中に赤玉 4 個と白玉 5 個が入っているとする。この袋から玉を 1 個取り出したとき、赤玉を取り出す事象を A、白玉を取り出す事象を B とする。取り出した玉は袋に戻さないとしておく。当然 $P(A) = \dfrac{4}{9}$ であり、$P(B) = \dfrac{5}{9}$ である。しかし、はじめに袋から赤玉が 1 個取り出された後では、後から白玉を取り出す確率は変化する。その場合は、袋の中には赤玉 3 個と白玉 5 個が入っている状態なので、赤玉が取られた後に白玉を取り出す確率は $P_A(B) = \dfrac{5}{8}$ である。

同じことを別の視点で見ることにしよう。赤玉が取られた後に白玉を取るということは、赤玉を取り出す事象 A と白玉を取り出す事象 B の共通部分 $A \cap B$ にあたる。その確率は $P(A \cap B)$ である。この場合の全事象であるが、はじめ袋には合計 9 個の玉があり、その後で赤玉が 1 個取られ合計 8 個になるのだから、玉の取り出し方は 9×8（通り）である。一方、赤玉が 1 個取られた後に

白玉を取る取り方は 4×5（通り）である。よって

$$P(A \cap B) = \frac{4 \times 5}{9 \times 8} = \frac{4}{9} \times \frac{5}{8} = P(A) \times P_A(B)$$

が成り立っているのだが、考えてみれば確率における乗法の性質が適用される自然な式になっている。

少し堅苦しい言い方ではあるが

$$\boxed{P(A \cap B) = P(A)P_A(B)}$$

を**従属事象の定理**と呼ぶことにしよう。

原因の確率

ある事象が起こった場合、この結果をもたらした原因である事象の確率を調べることができる。いま、互いに排反する事象 A、B があったとする。このとき A、B を原因とする事象 E が起こったとして、その確率を $P(E)$ とする。このとき、E が起こったことを条件にして、その原因を A、B どちらに求めるのが妥当か、言い換えれば、事象 A が原因で起こった確率を調べることは、$P_E(A)$ を求めることに他ならないのである。

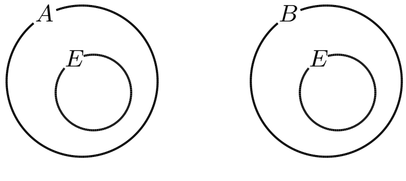
さて、ここで $P(E \cap A)$ を考えると、$P(E \cap A) = P(E)P_E(A)$ より、$P_E(A)$ について整理した上で、従属事象の定理を当てはめると

$$P_E(A) = \frac{P(E \cap A)}{P(E)} = \frac{P(A \cap E)}{P(E)} = \frac{P(A)P_A(E)}{P(E)} \quad (※)$$

である。また、事象 E が起こったのは、排反事象 A、B を原因とすることから、$A \cap E$ と $B \cap E$ の和を考えて

$$P(E) = P(A \cap E) + P(B \cap E) = P(A)P_A(E) + P(B)P_B(E) \quad (☆)$$

が成り立つ。このことは少し分かりにくいかもしれないので、図で説明しておこう。



A、B は排反だから A、B に共通部分はない。しかし A か B が原因で E が起こるのなら、E は A や B に含まれているはずである。A に含まれる E は $A \cap E$、B に含まれる E は $B \cap E$ なので、E が起こることは $(A \cap E) \cup (B \cap E)$ ということになる。このことを指して、$P(E) = P(A \cap E) + P(B \cap E)$ を計算したのである。

計算の結果、（☆）を（※）の分母へ代入すると

$$\boxed{P_E(A) = \frac{P(A)P_A(E)}{P(A)P_A(E) + P(B)P_B(E)}}$$

という関係式を導くことができるのだが、この式もまた分かりにくいかもしれない。

そこで、具体的な例を示すことにしよう。赤玉 1 個と白玉 4 個が入った袋 M と、赤玉 2 個と白玉 4 個が入った袋 N がある。袋 M か袋 N のどちらかから玉を 1 個取ったら赤玉であった。赤玉が袋 M から取り出された確率を求めてみよう。

まず、袋 M と袋 N を選ぶ確率は等しいと仮定しておくと $P(M) = \dfrac{1}{2}$、$P(N) = \dfrac{1}{2}$ である。袋 M を選べば赤玉を取る確率は $P_M(\text{赤}) = \dfrac{1}{5}$ で、袋 N を選べば赤玉を取る確率は $P_N(\text{赤}) = \dfrac{1}{3}$ である。したがって

$$P_{\text{赤}}(M) = \frac{P(M)P_M(\text{赤})}{P(M)P_M(\text{赤}) + P(N)P_N(\text{赤})} = \frac{\frac{1}{2} \cdot \frac{1}{5}}{\frac{1}{2} \cdot \frac{1}{5} + \frac{1}{2} \cdot \frac{1}{3}} = \frac{3}{8}$$

となる。この結果 $P_{\text{赤}}(N) = \dfrac{5}{8}$ も分かる。つまり赤玉の出所は、袋 M からより袋 N からである可能性の方が 3 : 5 の割合で高いことになる。

*　　*　　*

大層な計算をして、当たり前の結果を得ただけのような気がするかもしれない。袋 N は明らかに袋 M に余分な赤玉 1 個を入れたものになっているので、取り出された赤玉は袋 N から出された可能性が高いに決まっている。しかし、可能性はどの程度高いかと問われれば答えに窮するのではないだろうか。それを正確に言うのがこの計算なのである。

Microsoft Excel を使って、原因の確率をシミュレートしてみよう。ワークシートには上の例と同じ条件でシミュレートしてある。

◇	A	B	C	D	E	F
1	袋 M →赤	袋 N →赤	取り出し	判定		
2	=1/5	=1/3	（※ **C2**）	（※ **D2**）		
3	（※ **A3**）	（※ **B3**）	↓下へコピーする	↓下へコピーする		
4			↓	↓		
5			↓	↓		
6			↓	↓		

※ セルの式

```
（C2）  =RAND()
（D2）  =IF(C2<$A$3,”M”,IF(AND(C2>0.5,C2<$B$3),”N”,””))
（A3）  =A2*0.5
（B3）  =0.5+B2*0.5
```

A2 セルと B2 セルには袋 M と袋 N から赤玉が取り出される確率を入力する。袋 M か袋 N を選ぶ確率は半々であったから、A3 セルと B3 セルでは 0.5 が掛けられている。とくに B3 セルの式に 0.5 が加えられて

いるのは、RAND 関数で 0 以上 1 未満の実数が生成されたとき、0.5 未満なら袋 M を、0.5 以上なら袋 N を意味するようにしたかったからである。E 列で赤玉が袋 M か袋 N のどちらから取り出されたかを判定しているので、D 列に表示される M と N の数を調べれば、約 $3:5$ であることが確認できるだろう。

A2 セルと B2 セルの値を変えれば別のシミュレーションになるので、実際の計算結果にほぼ合うことが分かるだろう。∎