統計数理

石川顕一

http://ishiken.free.fr/lecture.html

10/18　組み合わせと確率
10/25　確率変数と確率分布
11/1　代表的な確率分布
11/8（前半）　ランダムウォークと破産問題
11/8（後半）　ブラウン運動と拡散
11/22　ノイズの理論

参考書
理工系の数学入門コース7・薩摩順吉著
「確率・統計」岩波書店

11/15は演習のみ

参考書

[1] 理工系の数学入門コース7・薩摩順吉著
「確率・統計」岩波書店 2001年

[2] コルモゴロフ、ジュルベンコ、プロホロフ 共著
丸山 哲郎、馬場 良和 共訳
「コルモゴロフの確率論入門」森北出版 2003年

[3] 岩波基礎物理シリーズ8・北原和夫著
「非平衡系の統計力学」岩波書店 1997年

統計数理

石川顕一

10/18　確率

- 順列と組み合わせ
- 直線上のランダムウォーク
- 確率の定義
- 確率の性質
- 条件付き確率

1－1　順列と組み合わせ

→ 参考書[1]

- アルファベット26文字が一つずつ書かれた26枚のカードを袋に入れ、そこから無作為に1枚ずつ取りだして、取りだした順に並べたら、BISになる確率は？
- A, F, I, I, M, N, N, O, O, R, Tの一字ずつを書いた11枚のカードを袋に入れ、そこから無作為に1枚ずつ取りだして、取りだした順に並べたら、INFORMATIONになる確率は？

⇒ 順列

- 順列(permutation)
  - 与えられた複数個のものからいくつかをとって、順番に1列に並べたものを順列という。

n 個の異なるものの中から、任意に r 個とって1列に並べる順列の数は

$$ {}_nP_r = n(n-1)(n-2)\cdots(n-r+1) = \frac{n!}{(n-r)!} $$

特に、異なる n 個のものを全部1列に並べる順列の数は

$$ {}_nP_n = n(n-1)(n-2)\cdots 2\cdot 1 = n! $$

[例]　${}_nP_r = {}_{n-1}P_r + r\,{}_{n-1}P_{r-1}$　⟹　n 個の場合と n -1個の場合を関係づける漸化式

- 重複順列
  - n 個の異なるものの中から、繰り返しを許して（同じものを何回使ってもよい）r個とり、1列に並べる順列（重複順列）の数は、n^r
  - [例] 1,2,3,4の4個の数字を用いて、3桁の自然数を作るとき、その総数は、積の法則より、4×4×4＝4^3＝64通り。

- 同じものがある場合の順列
  - [例] A, F, I, I, M, N, N, O, O, R, Tの一字ずつを書いた11枚のカードを袋に入れ、そこから無作為に1枚ずつ取りだし、取りだした順に並べてできる単語の数は？（辞書にある単語かどうかは気にしない）

n 個のもの c 個の組に分けられていて、同じ組に属するもの同士は区別できないが、異なる組に属するものは区別できるとき、これら n 個すべてを1列に並べる順列の数は、

$$\frac{n!}{n_1!n_2!\cdots n_c!} \qquad (n_1+n_2+\cdots+n_c=n)$$

各組に一つずつしかない時は、普通の順列になる。

- 組み合わせ(combination)
  - 与えられた複数個のものから、順序づけはしないでいくつか選んだ組を、組み合わせという。

n 個の異なるものから任意にr個とった組み合わせの数は

$$ {}_nC_r = \frac{{}_nP_r}{r!} = \frac{n(n-1)\mathcal{L}\,(n-r+1)}{r(r-1)\mathcal{L}\,2\cdot 1} = \frac{n!}{r!(n-r)!} $$

$\begin{pmatrix} n \\ r \end{pmatrix}$ と書くことも多い。

[例] 色の異なる5つの球のうち3つを選んで1つの組を作るときの組み合わせの数は、
（５×４×３）÷（３×２×１）＝１０
通りある。

[例]　${}_nC_r = {}_{n-1}C_r + {}_{n-1}C_{r-1}$ $\Longrightarrow$ n 個の場合とn -1個の場合を関係づける漸化式

$$ {}_2C_r = \mathcal{L} $$

$$ {}_3C_r = \mathcal{L} $$

$$ {}_4C_r = \mathcal{L} $$

- ２項定理(binomial theorem)
  - $(a+b)^n$の展開を示す公式

$$(a+b)^2 = a^2 + 2ab + b^2$$
$$(a+b)^3 = a^3 + 3a^2b + 3ab^2 + b^3$$

$$(a+b)^n = \underbrace{(a+b)(a+b)\cdots(a+b)}_{n\text{個}}$$

$a^{n-r}b^r$の係数は、n個の因子$(a+b)$から、bをr個選ぶ組み合わせ　${}_nC_r$

nが正の整数のとき、

$$(a+b)^n = \sum_{r=0}^{n} {}_nC_r a^{n-r} b^r$$

ニュートンの２項式

この結果から、　${}_nC_r = \binom{n}{r}$　を２項係数ともいう。

[例] $(5x\text{-}3y)^8$の展開式のx^3y^5の係数は

- 重複組み合わせ

[例] 2種類（赤白）のワインを売っている店で、3本のワインを買うとすれば、赤3、赤2白1、赤1白2、白3の4通りの買い方がある。

⇒ 4つの場所の1つに「しきり」を入れることに対応。

[例] ○○｜○＝赤2白1

n 個の異なるものから、繰り返しを許して r 個とるときの組み合わせの数は

$$ {}_{n}H_{r} = {}_{n+r-1}C_{r} = \frac{n(n+1)\cdots(n+r-1)}{r!} $$

- いくつかの組に分ける場合の組み合わせ

[例] 7人の学生を3人と4人の2つの組に分ける。

⇒ 7人をならべて、前の3人と後ろの4人に分ければいい。3人と4人の順番は問わないから7!/(3!×4!)=35通り

n 個の異なるをn_1個, n_2個, …, n_c個のc組に分ける組み合わせの数は

$$ \frac{n!}{n_1!n_2!\cdots n_c!} \qquad (n_1+n_2+\cdots+n_c=n) $$

「同じものがある場合の順列」と同じ

- 多項定理

いくつかの組に分ける場合の組み合わせ

$$(a_1+a_2+\cdots+a_m)^n=\sum\frac{n!}{n_1!n_2!\cdots n_m!}a_1^{n_1}a_2^{n_2}\cdots a_m^{n_m}$$

ただし和は、 $n_1\ge 0, n_2\ge 0, \cdots, n_m\ge 0$ で $n_1+n_2+\cdots+n_m=n$

を満たすすべての $n_1, n_2, \cdots, n_m$ についてとる。

[例] $(x-5y+3z)^6$ の展開式の xy^2z^3 の係数は、

1−2 直線上のランダムウォーク

→ 参考書[2]

確率論の応用範囲は、サイコロやトランプに関する問題だけではない！

- 気体の運動理論
- 拡散現象
- 雑音(ノイズ)
- 株価・為替
- ...

非平衡系の統計力学

数理ファイナンス

確率過程

個々の粒子の無秩序で雑然とした運動 ⇔ 全体として見た場合の明確で簡単な法則性

確率論(確率過程の理論)

・ブラウン運動

植物学者R. ブラウン 1827年

水に浮かんでいる花粉の粒子は、
たえず無秩序な運動をしていること
を発見。

花粉の生命力？

著作権処理の都合で、
この場所に挿入されていた
図を省略させて頂きます。

（J. ペランの実験結果による。出典：参考書[2] p.7）

すべての十分に細かい粒子の一般的性質であることが判明

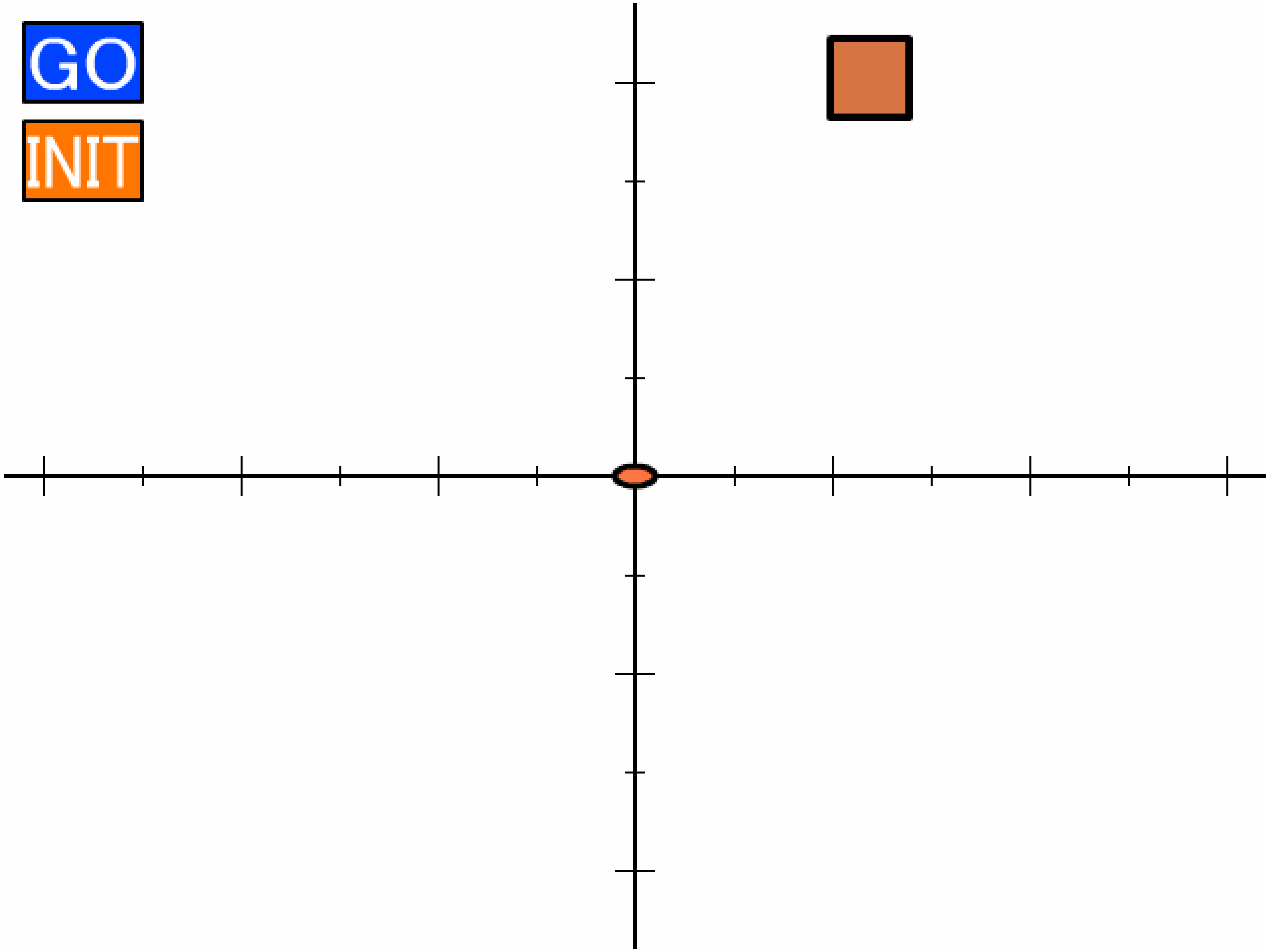
[例] 水槽中に落とした一滴のインクの拡散

$$r = a\sqrt{t}$$

インク滴の半径は時間に
比例しない。

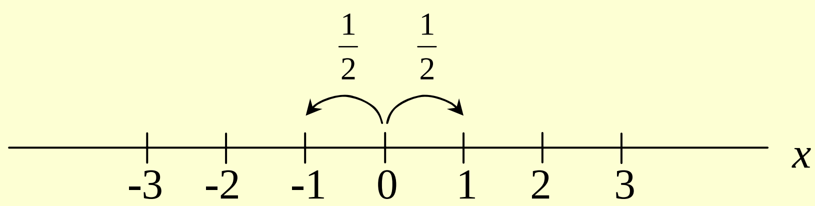
インク滴
の半径

・平面上のブラウン運動（バクテリアの運動のシミュレーション）

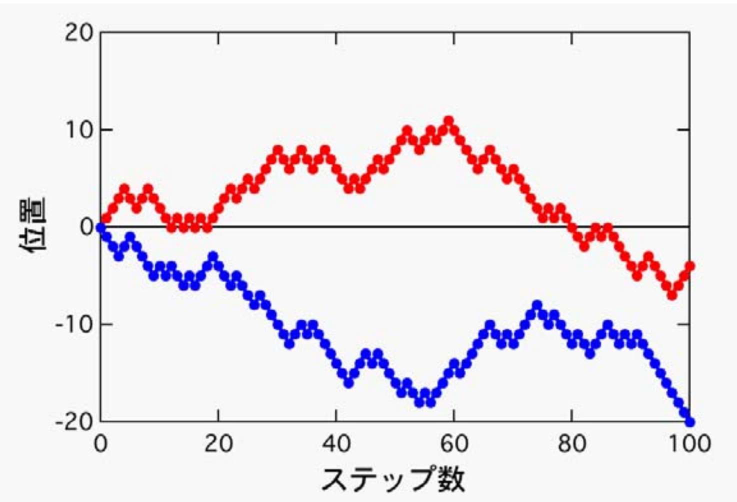


BIS卒業生岸勇気君作

・直線上のランダムウォーク



- 時刻t=0にm=0を出発
- 1回のジャンプごとに1だけ、右または左へ移動する。
- どの位置にいても次に右へ進む確率と左へ進む確率は1/2ずつ。



N ステップ後の粒子の位置 *m*(*N*)

$$m(0) = 0$$

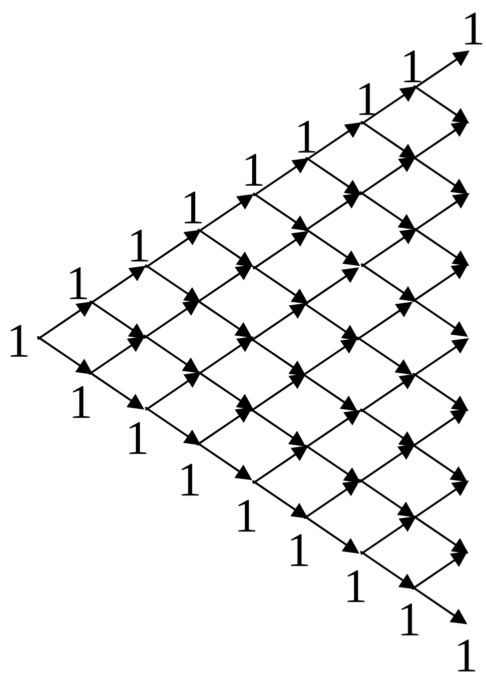
$$m(1) = -1,\ 1$$

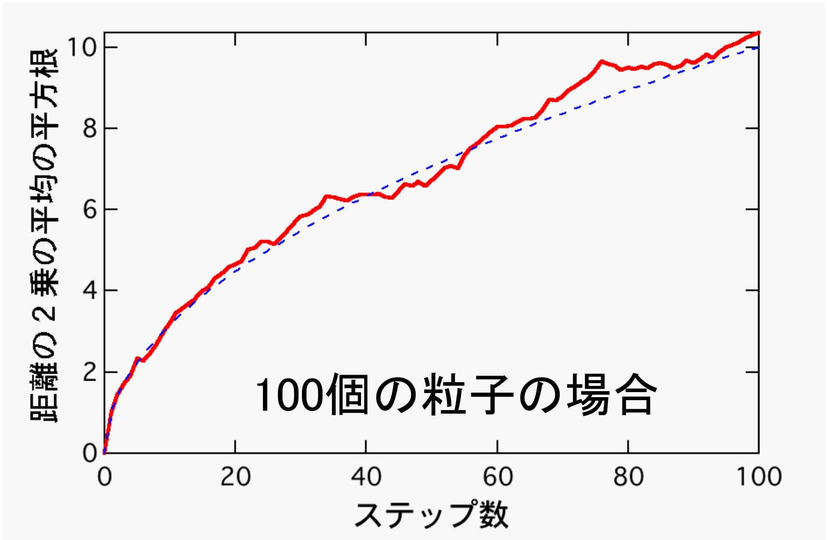
$$m(2) = -2,\ 0,\ 2$$

$$m(3) = -3,\ -1,\ 1,\ 3$$

$$m(N) = -N, -N+2, -N+4, \ldots, N-4, N-2, N = N-2k\ (k = 0, 1, \ldots, N)$$

・直線上の軌道数の計算





2項定理より　$\sum_{k=0}^{n} {}_nC_k = 2^n$　$\sum_{k=0}^{n} k\,{}_nC_k = n2^{n-1}$　$\sum_{k=0}^{n} k^2\,{}_nC_k = n(n+1)2^{n-2}$

これらを用いて、距離の2乗の平均の平方根は…

1−3 確率の定義

→ 参考書[1]

- 確率・統計で扱う対象
  - サイコロ振り、コイン投げ、電気的雑音の電圧測定…同じとみなされる条件のもとで、何回でも繰り返しのできること
  - 同じ大きさのたくさんの玉、容器中の気体分子…質の同じ個体が多数集まっている集団
- 試行(trial)：サイコロを振って目を読む、気体分子の運動エネルギーを測定する、などの操作
- 事象(event)：試行を行って得られる結果
  - [例] 3の目が出る、奇数目が出る
- 標本空間(sample space)：起こりうる結果の全体
  - 根元（こんげん）事象：それ以上にわけられない事象（[例] 3の目が出る）
  - 結合事象：2つ以上の根元事象を含む事象（[例] 奇数目が出る）

これらの事象では、1回1回の試行によってどの事象がえられるかは不確定である。

↓

回数を増やせば、ある規則性が存在する。

⇒ 確率の理論的考察

- 数学的確率
  - ラプラス(Laplace)によって与えられた。

> ある試行について、標本空間の大きさが n で、どの根元事象も同程度に確からしく起こるとする。標本空間の中で、ある事象 E をとり、E の起こる場合の数が r であるとき、E の確率 $P(E)$ を
>
> $$P(E)=\frac{r}{n}$$
>
> と定義する。（参考書[1] p.19）

［例］ 10枚の百円玉を投げて、6枚が表、4枚が裏となる確率はいくらか。ただし、表が出るのも裏が出るのも同様に確からしいとする。

［例］ 百円玉を10回繰り返して投げ、表なら○を、裏なら×をノートに記録する。

(A)　○○○○○○○○○○

(B)　○××○○×○○×○

のどちらの出方の方がどれくらい確率が高いか。

- 経験的確率（統計的確率）
  - 野球の打率、天気予報の当たる確率、不完全なサイコロの目の出方…

n 回試行を行った結果、ある事象 E が r 回起こったとする。n を大きくしていくとき、r/n が一定の値 p に近づくならば、E の確率 $P(E)$ を

$$P(E) = p = \lim_{n\to\infty} \frac{r}{n}$$

とする。（参考書[1] p.21）

［例］ 打率が.333のバッターが、ある試合の第1，第2打席でともに凡退した。第3打席でヒットを打つ確率はいくらか？（参考書[1] p.21）

［例］ （理想的な）サイコロを何回か続けて振ったところ、5回続けて6の目が出た。次に振ったときに、6の目が出る確率と、1の目が出る確率はどちらが高いか？

［例］ ある学科に40人の学生がいる。その中で誰かと誰かの誕生日が一致する確率はいくらか？

1―4　確率の性質

→ 参考書[1]

集合の概念を用いる。

- 標本空間をSとすると、Sは1つの集合であり、事象EはSの部分集合である。　$E \subset S$
- AとBの積事象：事象AとBが同時に起こる事象　$A \cap B$
- AとBの和集合：事象AとBの少なくとも一方がおこる事象　$A \cup B$
- Eの余事象：Sの中で、Eの起こらない事象　$\overline{E}$
- 空事象 ϕ：決して起こらない事象
- 排反(exclusive)：事象AとBが同時にはおこらないとき、AとBは互いに排反であるという。　$A \cap B = \phi$
  - [例] サイコロ振りで、Aを偶数目、Bを5の目とすると、AとBは互いに排反。
  - 根元事象はすべて互いに排反である。

- 確率の公理　　→ 参考書[1] p.23
  - 標本空間Sの各事象Eの確率$P(E)$は、以下の3つの条件を満たす。
  - 標本空間Sの各事象Eに対して、次の3つの条件を満たす実数$P(E)$が存在するとき、$P(E)$を事象Eが起こる確率という。

(1)　$0 \leq P(E) \leq 1$

(2)　$P(S) = 1,\ P(\phi) = 0$

(3)　$E_1, E_2, E_3,$ ・・が互いに排反な事象のとき、

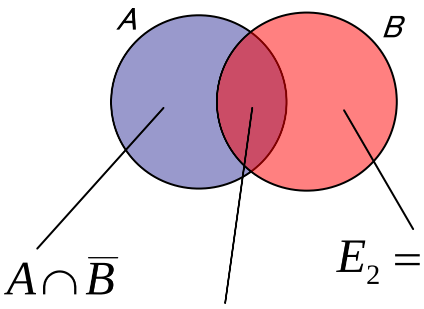
$$P(E_1 \cup E_2 \cup E_3 \cup \cdots) = P(E_1) + P(E_2) + P(E_3) + \cdots$$

事象の個数が無限でもいい。

- 確率の公理から導かれるいくつかの公式　　→ 参考書[1] p.24

加法公式

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$



$E_1 = A \cap \overline{B}$　$E_2 = \overline{A} \cap B$　$E_3 = A \cap B$

[例] よく切ったトランプ52枚（ジョーカーを含まない）から、1枚とりだして、そのカードがスペードである（A）か、または絵札である（B）確率は？

$$P(A) = \frac{13}{52},\ P(B) = \frac{12}{52},\ P(A \cap B) = \frac{3}{52} \Rightarrow P(A \cup B) = \frac{13}{52} + \frac{12}{52} - \frac{3}{52} = \frac{11}{26}$$

$$P(\overline{E}) = 1 - P(E) \quad \Longleftarrow \quad E \cup \overline{E} = S,\ E \cap \overline{E} = \phi$$

[例] 雨が降る確率が70%のとき、雨が降らない確率は1−0.7=0.3、すなわち30%である。

1−5　条件付き確率

- 条件付き確率　　→ 参考書[1] p.27

２つの事象A, Bがあって、Aが起こったという条件のもとでBが起こるという事象を$B|A$で表す。また、その確率$P(B|A)$を、条件AのもとでのBの条件付き確率(conditional probability)といい、

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

で定義する。

[例] トランプから１枚取り出す場合

スペードでなかったときのことは考えない。

A : スペードである事象

B : 絵札である事象

$B|A$: スペードであったときに、それが絵札である事象

$$P(B\,|\,A)=\frac{3}{13}\left(=\frac{3/52}{13/52}=\frac{P(A\cap B)}{P(A)}\right)$$

1−5　条件付き確率

- **乗法定理**　　→ 参考書[1] p.27

条件AのもとでのBの条件付き確率　$P(B|A)=\dfrac{P(A\cap B)}{P(A)}$

条件BのもとでのAの条件付き確率　$P(A|B)=\dfrac{P(A\cap B)}{P(B)}$

乗法定理　$P(A\cap B)=P(A)P(B|A)=P(B)P(A|B)$

[例]　くじ引き　　→ 参考書[1] p.27

10本のくじがあるとき、最初に引いた人が当たる事象をA、2番目に引いた人が当たる事象をBとする。

当たりくじが1本の場合

$$P(A)=\frac{1}{10}\qquad P(B)=P(A\cap B)+P(\bar{A}\cap B)=0+P(\bar{A})P(B|\bar{A})=\frac{9}{10}\cdot\frac{3}{9}=\frac{3}{10}$$

当たりくじが2本の場合

$$P(A)=\frac{2}{10}=\frac{1}{5}\qquad P(B)=P(A\cap B)+P(\bar{A}\cap B)=P(A)P(B|A)+P(\bar{A})P(B|\bar{A})$$

$$=\frac{2}{10}\cdot\frac{1}{9}+\frac{8}{10}\cdot\frac{2}{9}=\frac{1}{5}$$

1−5　条件付き確率

- ベイズの定理(Thomas Bayes)

[例] ある薬物検査は、ステロイド剤を使用している人に対して98%の確率で陽性を示し、また、使用していない人に対しても10%の確率で陽性を示す。あるサッカークラブでは部員の20%がステロイド剤を使用しているが、いま、部員の一人を検査したところ陽性であった。この部員がステロイド剤を使用している確率はいくらか。

元は何であったか or 原因の確率 or 事後確率

→ 遺伝子研究

スパムメール検出
（ベイジアンフィルター）

直観的に…

乗法定理をもちいて…

ベイズの定理　ある結果Eが、n個の互いに排反ですべての場合を尽くす原因$A_1, A_2, \cdot\cdot, A_n$によっているとき、そのうちの1つの$A_i$が原因である確率$P(A_i|E)$は

$$P(A_i \mid E) = \frac{P(A_i)P(E \mid A_i)}{P(A_1)P(E \mid A_1) + P(A_2)P(E \mid A_2) + \cdots + P(A_n)P(E \mid A_n)}$$

→ 参考書[1] p.30

1−5 条件付き確率

- ベイズの定理　　→ 参考書[1] p.30

ベイズの定理　ある結果Eが、n個の互いに排反ですべての場合を尽くす原因A_1, A_2, ・・, A_nによっているとき、そのうちの1つのA_iが原因である確率$P(A_i|E)$は

$$P(A_i\,|\,E)=\frac{P(A_i)P(E\,|\,A_i)}{P(A_1)P(E\,|\,A_1)+P(A_2)P(E\,|\,A_2)+\cdots+P(A_n)P(E\,|\,A_n)}$$

[例] 3つの機械A, B, Cのうち、生産量の10%をA, 30%をB, 60%をCが占めているとする。また、不良品の出る(E)割合が、Aは3%, Bは2%, Cは1%であるとする。1つの製品を取り出したところ不良品であったとき、それがAの製品である確率は、

$$P(A\,|\,E)=\frac{P(A)P(E\,|\,A)}{P(A)P(E\,|\,A)+P(B)P(E\,|\,B)+P(C)P(E\,|\,C)}$$

$$=\frac{0.1\times0.03}{0.1\times0.03+0.3\times0.02+0.6\times0.01}=20\%$$

→ 参考書[1] p.31

1−5　条件付き確率

- 統計的独立

AとBは統計的に独立　　$P(A \mid B) = P(A)$

乗法定理

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

一般に、n個の事象$A_1, A_2, \cdot\cdot, A_n$があるとき、それからとりだした任意個の事象$A_{i1}, A_{i2}, \cdot\cdot, A_{ik}$に対して、　　$(2 \le k \le n)$

$$P(A_{i_1} \cap A_{i_2} \cap \mathcal{L} \cap A_{i_k}) = P(A_{i_1})P(A_{i_2})\mathcal{L}\, P(A_{i_k})$$

が成り立つとき、事象$A_1, A_2, \cdot\cdot, A_n$は互いに統計的に独立。

[例] 3つの機械A, B, Cのうち、生産量の10%をA, 30%をB, 60%をCが占めているとする。また、不良品の出る(E)割合が、Aは8%, Bは2%, Cは1%であるとする。1つの製品を取り出したところ不良品であったとき、それがBの製品である確率は、

$$P(B \mid E) = \frac{0.3 \times 0.02}{0.1 \times 0.08 + 0.3 \times 0.02 + 0.6 \times 0.01} = 30\%$$

また　　$P(B) = 30\%$

⇒ BとEは統計的に独立　注意！AやCはEと統計的に独立でない。