

読んで役立つ・数学雑学シリーズ



ピタリとわかる 確率・統計解析入門

幸良井 幸美井 共著



誠文堂新光社

注) Microsoft Excelは米国Microsoft Corporationの米国およびその他の国における登録商標です。
また、Windowsの正式名称はMicrosoft Windows Operating Systemです。

まえがき

わが国では、多くの人々が確率・統計に関する教育をほとんど受けることなく大人になってしまいます。このような状況が長い間続いて現在に至っています。多分、確率や統計を学ぶ人は、特別な専門家だけで十分だと、考えられていたためかもしれません。

ところが、このところのコンピュータの発達とともに事情は変わってきました。つまり、情報化社会の到来です。この結果、現代人は広大な情報の海を自分で上手に泳がなければならなくなっていました。少ない退職金や年金を、株式などで上手に運用していかないと老後の生活が立ち行かなくなってきました。得られたデータから適切な情報を読み取ることができないと会社はつぶれてしまうのです。このようなとき、時代を泳ぐ強力な武器になるのが「確率・統計」の考え方なのです。確率・統計を過去に学校で教えてくれなかったからといいわけしても事態は改善されません。ゼロに近かった確率・統計の教養を、ある程度使えるレベルにまで高めなければならないのです。そしてコンピュータを操り、いろいろな資料からそれらの裏に潜む価値ある情報を探りあてねばなりません。これらのことを普通の人々が行わなければならない時代になってしまったのです。

このような時代背景を考えて、本書は確率・統計の初学者の学習の一助になればと、次の内容で編集しました。

「確率・統計の基本的な考え方がわかる」

「知っている和生活が楽しくなる」

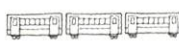
難しい数式は避けて、やさしい数学で解説してありますが、式が難しく感じる場合には文章だけ、もしくは図だけを見ればほぼその内容はわかるようになっています。読む順序はどのセクションからでもかまいませんが、一通り目を通していただけると幸いです。

なお本書に関連し、コンピュータシミュレーションについては拙著「パソコンで遊ぶ数学実験」(講談社)、表やグラフについては「Excel で学ぶ統計解析」(ナツメ社)を参考にしました。

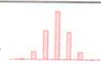
2005 年 1 月 著者

§ 1	「確定法則」と「確率法則」			8
§ 2	これだけ知っていれば確率は楽しめる			10
§ 3	場合の数を数えるための2つの知恵			12
§ 4	ファクトリアルが威力を発揮			14
§ 5	確率でよく使う「順列」と「組合せ」			16
§ 6	同じものを選んだ順列や組合せ			18
§ 7	「または」ときたら少し大変			20
§ 8	「少なくとも」ときたら意外に簡単			22
§ 9	集合を使って考えよう			24
§ 10	確率で使う集合同士の掛け算			26
§ 11	あちら立てればこちらが立たない			28
§ 12	影響するかしらないかを考慮すると			30
§ 13	よく間違える排反と独立			32
§ 14	(場合の数)/(場合の数)を広げると			34
§ 15	「表」や「裏」のままではつらいのです			36
§ 16	繰り返し起こる確率現象は			38
§ 17	数学なのに経験が大事			40
§ 18	あり地獄のような「大数の法則」			42
§ 19	独立にもいろいろ			44
§ 20	気分に影響された確率			46
§ 21	確率論のルーツは貴族の賭事			48
§ 22	高名な数学者でも間違っ			50
§ 23	見方で確率が変わってしまう			52
§ 24	原因を探る確率			54
§ 25	過去が未来に影響するマルコフ過程			56

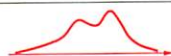


§ 26	期待値を知らずして賭をすべからず		58
§ 27	有限な人生に不公平は当たり前		60
§ 28	貧乏人が賭で金持ちに勝つ方法		62
§ 29	勝率が5割を少し超せば大金持ち		64
§ 30	「あみだくじ」に落とし穴		66
§ 31	「多勢に無勢」というけれど		68
§ 32	○×試験は素敵な試験		70
§ 33	下手な鉄砲も数撃ちゃ当たる		72
§ 34	くじ引きは先を争うことなかれ		74
§ 35	家族計画で男女比は変わらない		76
§ 36	意外に多い同じ日の誕生日		78
§ 37	サイコロの起源は占いの道具		80
§ 38	ジャンケンは少人数で		82
§ 39	針を落とせば円周率		84
§ 40	乗車車両は何両目か		86
§ 41	封筒の宛名と中身が違う		88
§ 42	酔っぱらいの行き着く先は		90
§ 43	野生動物の数は		92
§ 44	ミリオン分の1が1億円を産む		94
§ 45	最短コースにいろいろな数学		96
§ 46	「丁か半か」というけれど		98
§ 47	伝言ゲームとマルコフ過程		100
§ 48	一口に「乱数」というけれど		102
§ 49	疑似乱数はこうやって作る		104
§ 50	一様乱数からいろいろな乱数		106

§ 51	確率の歴史	108
§ 52	データの山は宝の山	110
§ 53	いくつに分類したらよいのか	112
§ 54	「一を聞いて十を知る」技	114
§ 55	どのくらい取り出せばよいのか	116
§ 56	もどすか、もどさないかの微妙な違い	118
§ 57	一口に平均というけれど	120
§ 58	分散は世界の多様性のあかし	122
§ 59	分布の特徴は平均と分散に集約	124
§ 60	3つの代表値の面白い関係	126
§ 61	分散は n 割、それとも $n-1$ 割	128
§ 62	人口密度があるように確率密度あり	130
§ 63	確率分布グラフで一目瞭然	132
§ 64	ジャスト 165cm である確率は 0	134
§ 65	サイコロの確率分布は二項分布	136
§ 66	何回目に初めて表が出るかは幾何分布	138
§ 67	希な現象はポアソン分布	140
§ 68	誤差の研究から正規分布	142
§ 69	「正規分布」作成マシーン	144
§ 70	標本分散の分布は χ^2 分布	146
§ 71	分散比の分布は F 分布	148
§ 72	標本平均の分布は t 分布	150
§ 73	分布の非対称性や尖り具合	152
§ 74	チェビシェフは万能	154
§ 75	変数なのに自由度	156



§ 76	標準化によって同じスケール	158
§ 77	偏差値も標準化の一種	160
§ 78	相関は原因結果ではありません	162
§ 79	相関関係はベクトルと絡む	164
§ 80	推定の妥当性には確率が	166
§ 81	点推定と区間推定、どっちが偉い	168
§ 82	資料が多いときの平均値の推定は	170
§ 83	資料が少ないときの平均値の推定は	172
§ 84	支持率の推定は	174
§ 85	分散も推定の対象に	176
§ 86	危険回避が検定の考え方	178
§ 87	検定そのものは機械的	180
§ 88	検定に2種類の誤り	182
§ 89	いいたいことが対立仮説	184
§ 90	ノンパラメトリックとは	186
§ 91	統計資料にはウラがある	188
§ 92	RDD 法はどんな抽出法	190
§ 93	このサイコロは正常かな？	192
§ 94	鯛焼き職人の技量のバラツキは違うか	194
§ 95	1000 年に一度は世論調査の対象に	196
§ 96	ポートフォリオで低リスク高収益率	198
§ 97	貯蓄現在高は 1688 万円	200
§ 98	パソコンは確率分布表を不要にした	202
§ 99	円周率とネイピアの数は非常に大事	204
§ 100	統計の歴史	206



ここが要点！

原因を与えれば結果が確定する法則を確定法則といいます。これに対して、原因を与えても結果がゆらいで確定しない法則を確率法則といいます。科学の法則は、大きくこの2つに大別できます。

やさしい解説

地球上で高いところからボールを落とせば、空中に静止することはありません。必ず地上に到達します。もし、空気抵抗を無視すれば、 t 秒間の落下距離 y は、

$$y = \frac{1}{2} g t^2$$

となります。

ただし、 g は重力加速度です。これが落下の法則です。これは、私ではなく彼女が落としても結果は同じです。もし、そうでなければ法則そのものが間違いとして意味を失います。

このように、同じ原因を与えれば結果が同じになる法則が確定法則なのです。中学や高等学校で学ぶ物理や化学の法則は、この確定法則にあてはまります。

これに対して、原因を与えても結果がピタッと確定しない法則があります。たとえば、男女の出生比はほぼ1:1であるという出生の法則がそうです。男女の出生については1:1が

期待されていますが、実際にそうなるわけではありません。

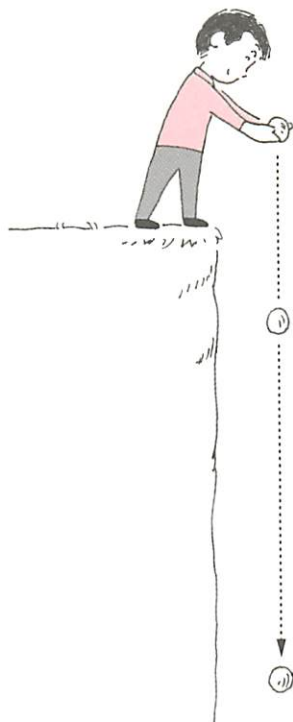
これはちょうどコインを投げて、その表裏に着目する場合と同様です。コインの表裏の出方は1:1が期待されていますが、実際には表が多かったり、また、その逆だったりします。表と裏がちょうど半分ずつ出るほうが希に感じるくらいです。

出生についても、地域や家庭において、男が多かったり女が多かったりします。しかし、全体としては出生比率はほぼ1:1なのです。この比が大きく変わることはありません。このように、個々の現象は結果が一通りにはありませんが、全体としてはある傾向に従う法則が確率法則ということになります。つまり、結果に確率が関与するのが確率法則なのです。

世の中には結果が一通りに決まる確定現象よりも、結果が一通りに決まらない不確定現象の方がはるかに多いのです。しかし、確率や統計的な手法が不確定現象を解明できるのは、まだ、それほど多くはなさそうです。

確定現象には確定法則が，不確定現象には確率法則が

1 ボールを放せば落ちる（確定現象）



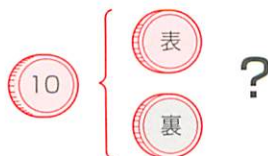
確定法則

$$y = \frac{1}{2}gt^2$$

y は落下距離
 t は落下時間
 g は重力加速度

2 10 円玉を投げれば表か裏が出る（不確定現象）

1回投げる



多数回投げる



確率法則

..... 表の出る相対度数はほぼ $\frac{1}{2}$

ここが要点!

確率現象は、多くの人が日常生活の中で体験しています。コインを投げて物事を決めたり、ジャンケンやあみだくじで役員を決めたりというようにです。このような確率現象は、きわめて初等的な数学の知識だけで解明し楽しむことができます。

注) もちろん、最先端の確率・統計理論では高度な数学が駆使されています。

やさしい解説

ある試行によって起こることが可能なすべての結果の総数が n であり、どの場合も同様な確からしさで起こるものとします。このうち、ある事柄 A の起こる場合の数が m ならば、その事柄 A の起こる確率は、

$$P(A) = \frac{\text{事象 } A \text{ の場合の数}}{\text{全体的場合の数}} = \frac{m}{n}$$

と考えます。

注) 確率は、英語で Probability。

つまり、「場合の数」分の「場合の数」をもって確率と考えるのです。このため、多くの人が気軽に確率を論じることができるのです。ただ、くれぐれも注意したいことは、「すべての結果の総数が n であり、**どの場合も同様な確からしさで起こるものとする**」という前提です。この前提を無視すると、とんでもない結論に達しますので注意してください。たとえば、「明日は陽が昇るか昇らないかの 2 通り。よって、

明日、陽が昇る確率は $1/2$ 」。「志望校に合格するかしないかの 2 通り。よって僕が志望校に合格する確率は $1/2$ 」, ……。という具合にです。

この「同様な確からしさ」をしっかり押さえれば、十分有意義な確率を求めることができます。

(例 1) サイコロを投げて偶数の目が出る確率は、

$$\frac{3}{6} = \frac{1}{2}$$

(例 2) 3人でジャンケンをしてあいこになる確率は、

$$\frac{3+6}{3 \times 3 \times 3} = \frac{9}{27} = \frac{1}{3}$$

(例 3) ジョーカーを除く 52 枚のトランプから 1 枚のカードを抜くとき、ハートである確率は、

$$\frac{13}{52} = \frac{1}{4}$$

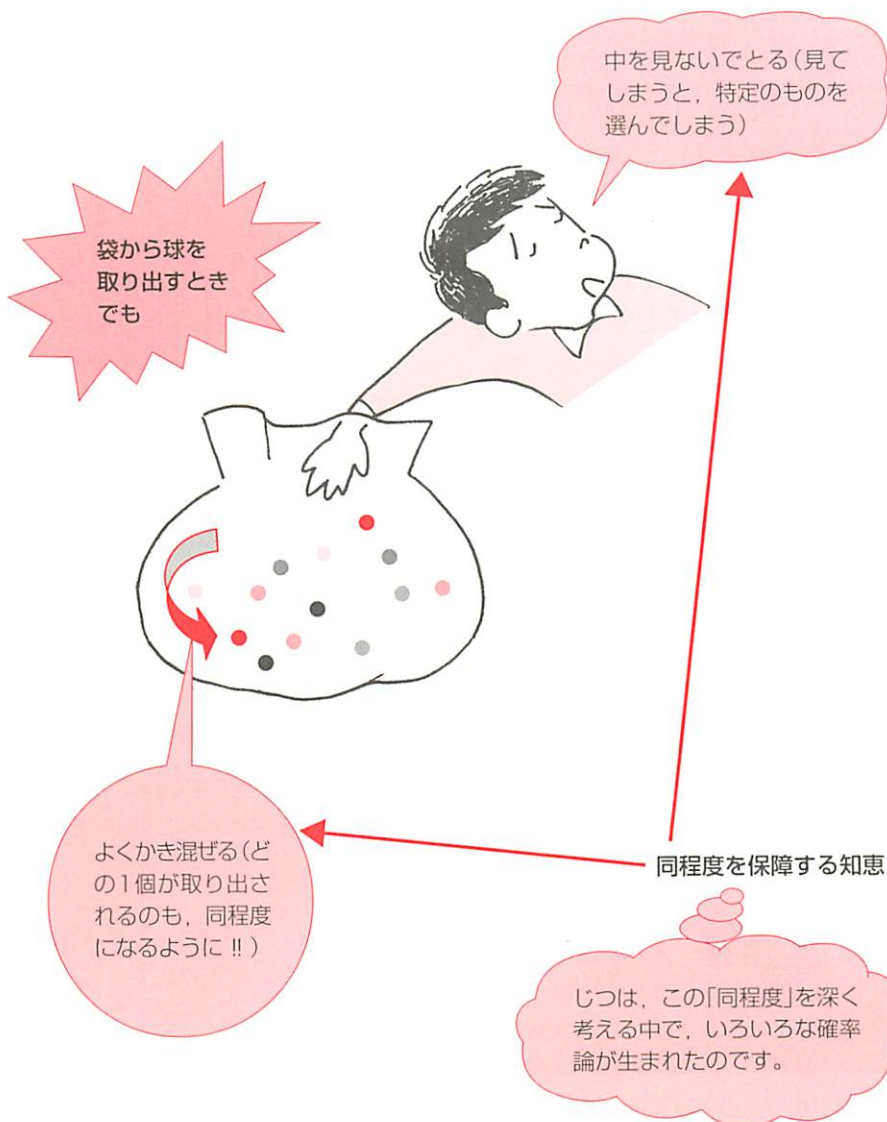
……

このほか、いろいろな確率が求められますから挑戦してみましょう。

場合の数さえ数えれば確率がわかる！

考えている場合の数
全体の場合の数 が求める確率

このとき、すべての場合が同程度で起こることが大前提 !!



ここが要点！

ある事柄の場合の数を正確に数えあげることは、確率を求める上で基本となります。しかし、複雑な事柄の場合の数となると、数えあげること自体が大変なことです。こんなとき役に立つ考え方が「和の法則」と「積の法則」です。どんな複雑な事柄でも解きほぐせば、この2つの考え方で解決できることが多いのです。

やさしい解説

和の法則とは、次の考え方です。

「2つの事柄AとBがあって、これらは同時に起こらないものとする。Aの起こり方が m 通りで、Bの起こり方が n 通りならば、AかBのいずれかが起こる起こり方は $m+n$ 通りである」

考え方そのものは常識的で難しくはありませんが、前提条件の「これらは同時に起こらないものとする」は押さえておいてください。

(例) 2つのサイコロ甲、乙を投げたとき、それらの目の和が5または7になる場合の数を求めてみましょう。目の出方は、次図のようになります。

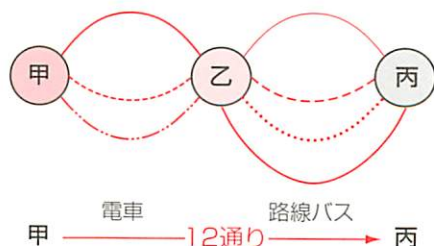
(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

ここでカッコ内の左は甲の目、右は乙の目です。この表から、目の和が5になるのは4通りで、目の和が7になるのは6通りで、これらは同時には起こりません。したがって、和の法則から答えは $4+6$ の10通りです。

積の法則とは次の考え方です。

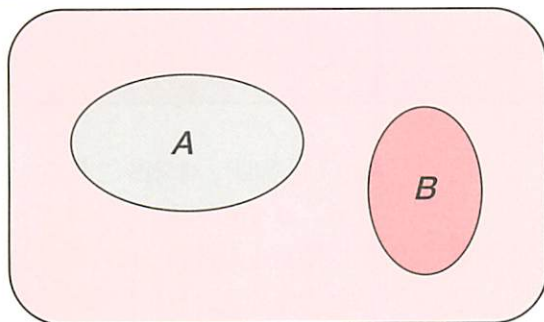
「2つの事柄AとBがあって、Aの起こり方は m 通りで、そのそれぞれに対してBの起こり方は n 通りとする。このときAの次にBが引き続いて起こる起こり方は $m \times n$ 通りである」

(例) 甲から乙には3種類の電車が、乙から丙には4種類の路線バスが運行しています。このとき、甲から乙を経て丙に乗り物で行く方法は、積の法則から 3×4 の12通りです。



足すから「和の法則」、掛けるから「積の法則」

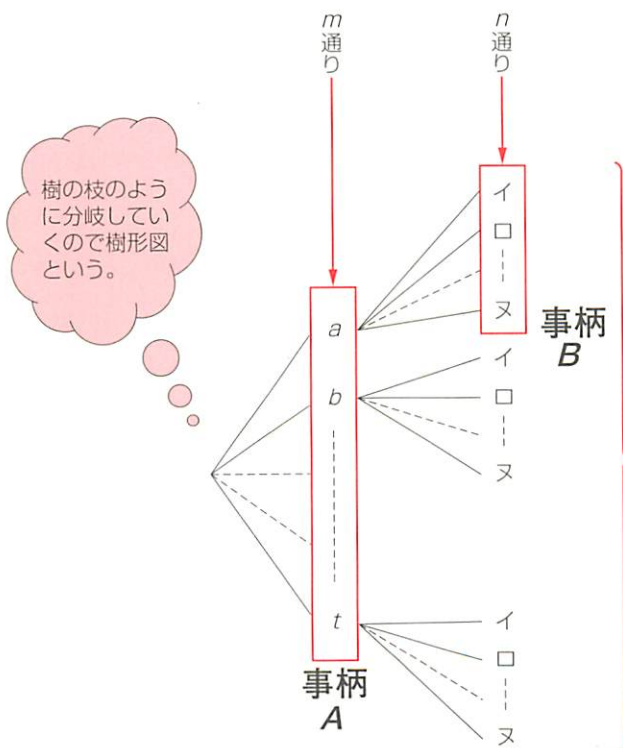
1 和の法則は集合で見るとわかりやすい



和の法則

集合AとBに共通部分
がなければ「Aまたは
B」の要素の個数は、
単純におのおのの要素
の個数を足せばよい。

2 積の法則は樹形図で見るとわかりやすい



積の法則

事柄Aに続いて、事柄Bが
起こるのは $m \times n$ 通り。

ここが要点!

ある場合の数を数え上げるときに便利で強力な記号があります。それが「 $n!$ 」で、 n の階乗 (Factorial) と読みます。

注) 階乗を一般化したものに、統計学で重要なガンマ関数があります。

やさしい解説

自然数 n に対して $n!$ とは、1 から n までの自然数をすべて掛け合わせたもので n の階乗と読みます。すなわち、

$$n! = n \times (n-1) \times (n-2) \times \cdots \times 3 \times 2 \times 1$$

のことです。この定義では $0!$ は定義されていませんが、他との整合性を考えて $0! = 1$ とします。

この階乗は、並べ方の数を数えるときによく使われます。たとえば、7 人が横1列に並ぶ並び方は、

$$7! = 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1 = 5040$$

通りです。7 人が手を取り合って輪になる並び方は、

$$(7-1)! = 6 \times 5 \times 4 \times 3 \times 2 \times 1 = 720$$

通りです。なお、後で紹介する順列 ${}_nP_r$ や組合せ ${}_nC_r$ もこの階乗をフルに使って定義されるものです。

それにしても、なぜ階乗に記号 $!$ が使われたのでしょうか。一説には n が大きくなると $n!$ はビックリするほど大きくなるからだといわれています

(右ページ参照)。本当かどうか?

それにしても、極めて大きな n に対して、階乗を定義通りに計算しようとすると大変です。もし、近似値でよければ、次のスターリング (Stirling) の公式を使うと便利です。

$$n! \doteq \sqrt{2\pi n} n^{n+\frac{1}{2}} e^{-n}$$

じつによく $n!$ を近似します (右ページ参照)。なお、 $n!$ の n は自然数ですが、この n を実数に拡張したものにオイラーの考えた次のガンマ (Gamma) 関数があります。

$$\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} dt \quad (x > 0)$$

ここで、 e はネイピアの数です。難しい関数ですね。理解する必要はありませんが、この関数は F 分布、 t 分布、 χ^2 分布など統計学上極めて大事な分布と密接な関係があることだけ気にとめておいてください。このガンマ関数は実数 λ に対して、

$\Gamma(\lambda) = (\lambda-1) \Gamma(\lambda-1) \quad (\lambda \geq 2)$ が成立し、特に x が自然数であれば、

$\Gamma(x) = (x-1)!$ が成り立ちます。

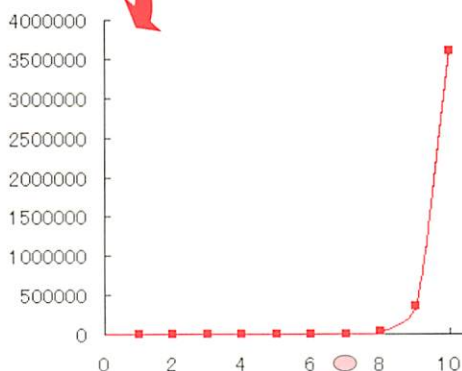
$n!$ は n が大きいと驚くほど大きい!

n	$n!$	スターリングの公式
1	1	0.9
2	2	1.9
3	6	5.8
4	24	23.5
5	120	118.0
6	720	710.1
7	5040	4980.4
8	40320	39902.4
9	362880	359536.9
10	3628800	3598695.6

95	1.033E+148	1.0321E+148
96	9.9168E+149	9.9082E+149
97	9.6193E+151	9.611E+151
98	9.4269E+153	9.4189E+153
99	9.3326E+155	9.3248E+155
100	9.3326E+157	9.3248E+157

スターリングの公式は
 $n!$ によく似合う。

グラフ化



似ているね!!

$n!$ のグラフとスターリング
の公式のグラフはほぼ重なっ
ている。

Memo 和の記号と積の記号

n 個の数 $\{a_1, a_2, a_3, \dots, a_n\}$ をすべて足
したものや、すべて掛け合わせたものは、

$$a_1 + a_2 + a_3 + \dots + a_n, \quad a_1 \times a_2 \times a_3 \times \dots \times a_n$$

などと表現されますが、なんだか締まりません。そこで、和と積に関しては特別な記号が用意されています。それが Σ (シグマ) と Π (パイ) です。これらは次のように定義されています。

$$\sum_{i=1}^n a_i = a_1 + a_2 + a_3 + \dots + a_n, \quad \prod_{i=1}^n a_i = a_1 \times a_2 \times a_3 \times \dots \times a_n$$

$$(\text{例}) \quad \sum_{k=1}^5 k = 1 + 2 + 3 + 4 + 5, \quad \prod_{i=1}^7 i = 1 \times 2 \times 3 \times 4 \times 5 \times 6 \times 7$$

ここが要点！

確率や統計では、「並べ方の総数」や「選び方の総数」を問題にすることがよくあります。そのため、これらの数を表現する記号が用意されています。それが、「 ${}_nP_r$ 」と「 ${}_nC_r$ 」です。これらの記号を上手に使うと、順列や組合せの問題が考えやすくなります。

注) 記号 ${}_nC_r$ と記号 $\binom{n}{r}$ は同じ意味で使われます。

やさしい解説

「相異なる n 個のものから r 個を選んで、1列に並べる仕方の総数のことを ${}_nP_r$ 」と書きます。

${}_nP_r$ は n から順次1ずつ引いた数を r 個掛け合わせたもので、次の計算で値を求めることができます。

$${}_nP_r = \underbrace{n(n-1)(n-2)\cdots(n-r+1)}_{n\text{個数の積}}$$

このことの成立理由は、1から r 番目までの指定席を想定し、そこに n 人から r 人選んで座らせると考えればわかります。つまり、最初の席には誰でもいいから n 通りの座り方があります。そのおのおの場合に対して2番目の席には、1番目に座った1人を除いた $n-1$ 通りの座り方があります。3番目の席には1、2番目のおのおのに対して2人を除いた $n-2$ 通りの座り方があります。このように考えていくと、 n から順次1ずつ引いた数を r 個

掛け合わせた

$$n(n-1)(n-2)\cdots(n-r+1)$$

であることがわかります。

次に組合せについて考えてみましょう。「相異なる n 個のものから r 個を選んでできる組合せの総数のことを ${}_nC_r$ 」と書きます。このは ${}_nC_r$ は、 ${}_nP_r$ を $r!$ で割ったものです。つまり、

$${}_nC_r = \frac{{}_nP_r}{r!}$$

ただし、

$$r! = r(r-1)(r-2)\cdots 3 \times 2 \times 1$$

この成立理由は、

$$\begin{aligned} {}_nP_r &= \text{相異なる } n \text{ 個のものから } r \text{ 個} \\ &\text{を選んで、そのおのおのを1} \\ &\text{列に並べる仕方の総数} \\ &= {}_nC_r \times r! \end{aligned}$$

からわかります。

(例)

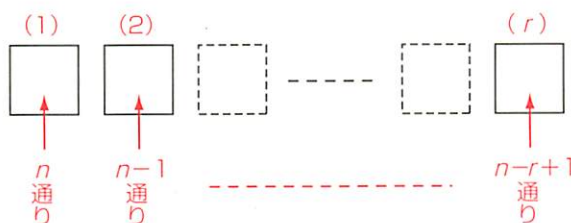
$${}_{10}P_3 = 10 \times 9 \times 8 = 720$$

$${}_{10}C_3 = \frac{{}_{10}P_3}{3!} = \frac{10 \times 9 \times 8}{3 \times 2 \times 1} = 120$$

nP_r , nC_r を使って場合の数を効率よく算出

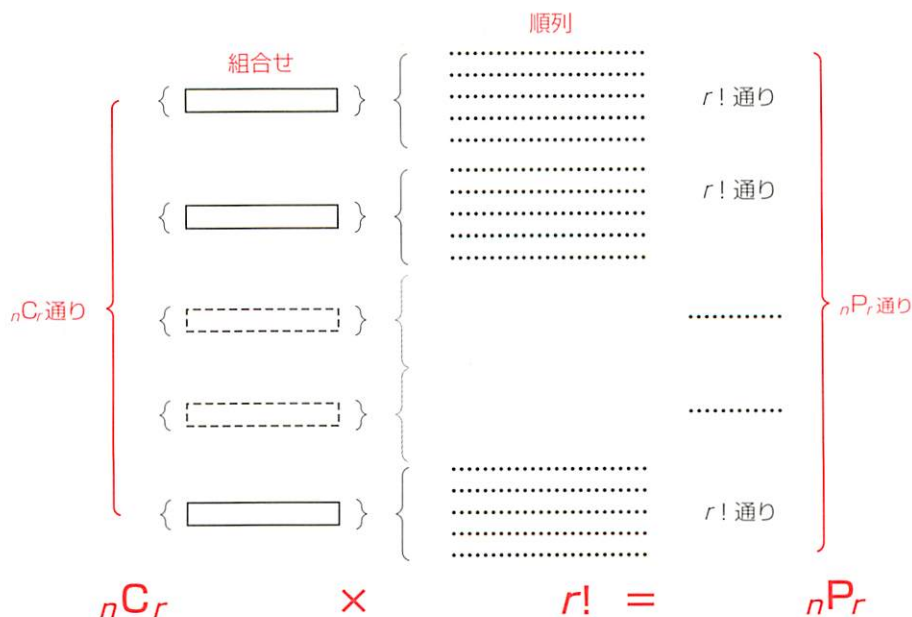
1 順列の計算

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



2 組合せの計算

$${}_nC_r = \frac{{}_nP_r}{r!} = \frac{n!}{(n-r)!r!}$$



ここが要点!

「順列」や「組合せ」において、同じものを選んで並べる場合や、組合せる場合があります。このような場合の数を数え上げるには、「重複順列 n^r 」や「重複組合せ ${}_nH_r$ 」の考え方を使うと便利です。

やさしい解説

ここに、1から5まで書かれた5枚のカードがあるとします。



これら、5枚のカードを横1列に並べて5桁の数を作ると、全部で、

${}_5P_5 = 5! = 5 \times 4 \times 3 \times 2 \times 1 = 120$ 通りできます。これが、通常の順列の考えです。ここで条件を変えて、これら5枚のカードを何回使ってもよいことにします。すると、5桁の数は「積の法則」から全部で、

$$5 \times 5 \times 5 \times 5 \times 5 = 3125$$

通りできます。万の位も千の位も百も十も一も、どの位にも5通りの数値が入り得るからです。一般に「相異なる n 個のものから、繰り返しとることを許して、 r 個にとって作った順列の数は n^r 」となります。これが**重複順列**の考えです。

次に「重複組合せ」を調べてみましょう。たとえば、3個の数字 $\{1, 2, 3\}$ から繰り返しとることを許して、2個選ぶ場合を考えてみることにします。

右の数値が左の数値以上になるようにして重複組合せを書き出すと、全部で6個あることがわかります。

$$\{1, 1\}, \{1, 2\}, \{1, 3\}, \{2, 2\}, \{2, 3\}, \{3, 3\} \dots\dots\dots \textcircled{1}$$

ここで、左の数に0、右の数に1を加えると式①は、

$$\{1, 2\}, \{1, 3\}, \{1, 4\}, \{2, 3\}, \{2, 4\}, \{3, 4\} \dots\dots\dots \textcircled{2}$$

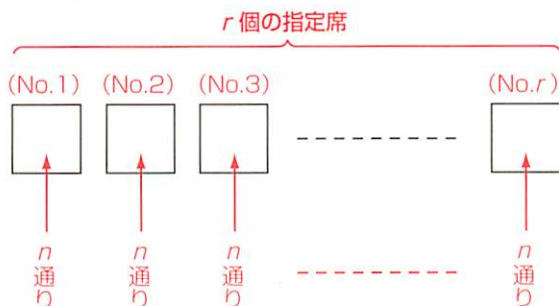
となります。式①と②では組合せの数は同じです。ここで、式②は異なる4個のものから2個選ぶ ${}_4C_2$ だけあります。したがって、 $\{1, 2, 3\}$ から重複を許して2個選ぶ選び方は ${}_4C_2$ 通りとなります。

一般に $\{1, 2, 3, \dots, n\}$ から r 個選ぶ場合は、右の数値が左の数値以上になるように書き出し、これに左から0, 1, 2, $\dots$, $(r-1)$ を加えることによって、異なる $n+r-1$ 個から r 個選ぶ組合せになります。

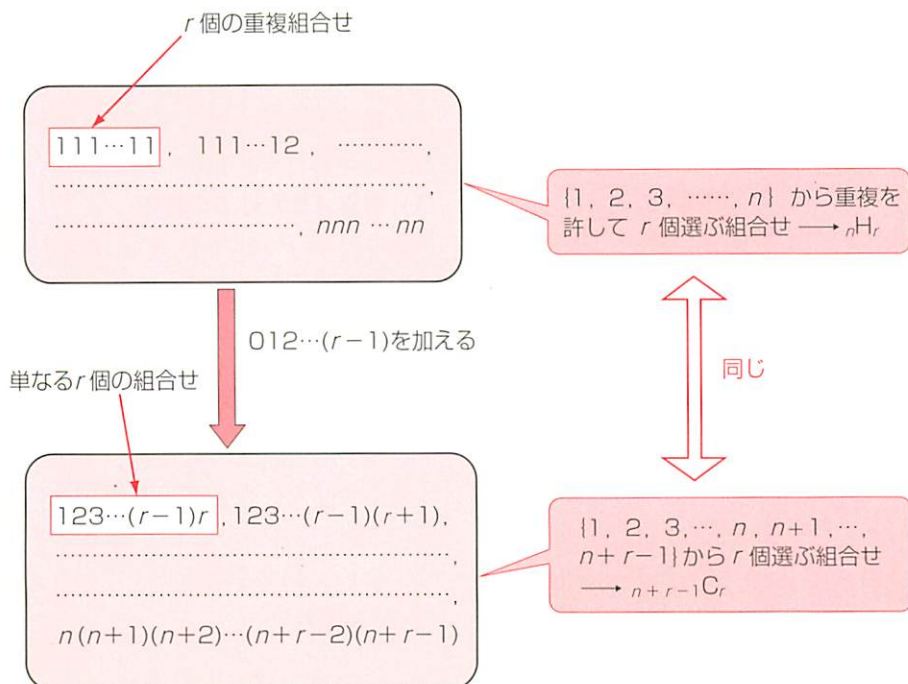
以上のことから、「相異なる n 個のものから、繰り返しとることを許して、 r 個にとって作った組み合わせの総数を ${}_nH_r$ で表すと、 ${}_nH_r = {}_{n+r-1}C_r$ 」となります。

重複順列・組合せは同じものを何回も使う

1 異なる n 個のものを重複を許して r 個並べる…… n^r



2 異なる n 個のものを重複を許して r 個選ぶ…… ${}_nH_r = {}_{n+r-1}C_r$



ここが要点!

「事象 A または B のいずれかが起こる確率を求めなさい」ということがよくあります。この「または」ときたら要注意です。事態は少し複雑だと考えた方がいいでしょう。

やさしい解説

たとえばここに、1から10までの10個の数値の集合

$$U = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$$

があります。ここから、1個の数値をとり出すとき、これが、2または3の倍数である確率を求めることにします。2の倍数である事象を A 、3の倍数である事象を B とすると、

$$A = \{2, 4, 6, 8, 10\}, B = \{3, 6, 9\}$$

となります。したがって、

$$P(A) = \frac{A \text{ の要素の数}}{U \text{ の要素の数}} = \frac{5}{10} = \frac{1}{2}$$

$$P(B) = \frac{B \text{ の要素の数}}{U \text{ の要素の数}} = \frac{3}{10}$$

すると、2または3の倍数である確率はこれらを加えて、

$$\frac{5}{10} + \frac{3}{10} = \frac{8}{10} = \frac{4}{5}$$

としたくなります。しかし、よく見ると事象 A と B の中に6が共通に含まれていることがわかります。したがって、この答えは少し多めに見積ったことに

なります。

事象 A または B は、

$$A \cup B = \{2, 3, 4, 6, 8, 9, 10\}$$

となりますから、正しい答えは、

$$P(A \cup B) = \frac{7}{10}$$

となります。

このように、「または」で表現された確率を求めるには絶えずダブリに気をつけねばなりません。そのためには、次の**確率の加法**を使って正しく処理する必要があります。

(1) 事象 A または B の確率

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

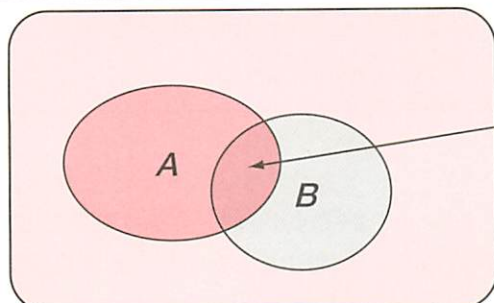
(2) 事象 A または B または C の確率

$$\begin{aligned} P(A \cup B \cup C) &= P(A) + P(B) + P(C) \\ &\quad - P(A \cap B) - P(B \cap C) \\ &\quad - P(C \cap A) + P(A \cap B \cap C) \end{aligned}$$

事象が4つ以上になると、もはや手に負えません。ただし、事象相互の関係に「排反」という条件があると事情は一変し、確率を求めやすくなります (§11 参照)。

「または」の確率計算はダブリに注意

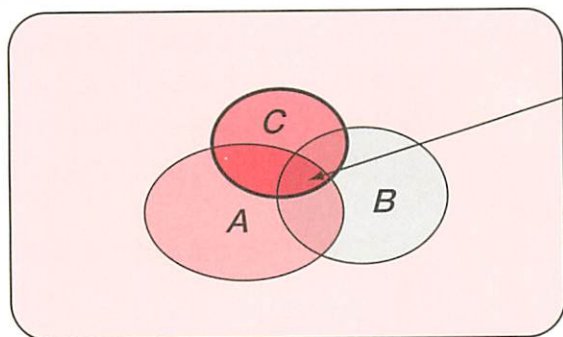
1 2つの事象AまたはBが起こる確率



このダブリに注意

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

2 3つの事象AまたはBまたはCが起こる確率



3つが重なっている、このダブリはとくに注意

$$\begin{aligned} P(A \cup B \cup C) \\ = P(A) + P(B) + P(C) - P(A \cap B) - P(B \cap C) \\ - P(C \cap A) + P(A \cap B \cap C) \end{aligned}$$

Memo 事象

サイコロを投げるというような試行の結果として起こる事柄を「事象」といいます。確率の世界では、「事象」は「集合」と同じです。

ここが要点！

「少なくとも1個が……」とか、「少なくとも1人が……」などというように、「少なくとも」という条件のついた確率現象は多くの場合、これを否定した世界を考えると簡単に解決できます。

やさしい解説

たとえば、黒球3個、白球4個入った袋からランダムに3個取り出すとき、少なくとも1個が黒球である確率を求めてみましょう。「少なくとも1個が黒球……」というわけですから言葉通りに現象を捉えて、次の3つの事象A, B, Cに分けて考えてみます。

A: 1個が黒で、2個が白

B: 2個が黒で、1個が白

C: 3個が黒で、0個が白

これらの確率を求めると、次のようになります。

$$P(A) = \frac{{}_3C_1 \times {}_4C_2}{{}_7C_3} = \frac{18}{35} \quad \cdots \cdots ①$$

$$P(B) = \frac{{}_3C_2 \times {}_4C_1}{{}_7C_3} = \frac{12}{35} \quad \cdots \cdots ②$$

$$P(C) = \frac{{}_3C_3 \times {}_4C_0}{{}_7C_3} = \frac{1}{35} \quad \cdots \cdots ③$$

事象A, B, Cは排反 (§11) なので「少なくとも1個が黒球である確率」は式①, ②, ③を単純に加えて、

$$\frac{18}{35} + \frac{12}{35} + \frac{1}{35} = \frac{31}{35}$$

となります。しかし、これはこれで正解ですが、何と冗長な解答でしょうか。単純明快をモットーとする数学では満足しかねます。

ちょっと考えてみると、「少なくとも1個が黒球である」ということを否定した事象は「黒球がない」こと、つまり、「3個とも白球」となります。そして、否定した事象同士は排反となり確率の和は1ですから、「少なくとも1個が黒球である確率」は1から「3個とも白球である確率」を引けばよいことになります。

ここで、「3個とも白球である確率」は、

$$\frac{{}_4C_3}{{}_7C_3} = \frac{4}{35}$$

ですから、求める確率は、

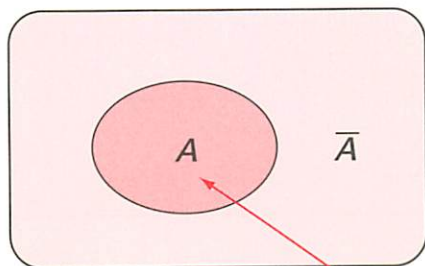
$$1 - \frac{4}{35} = \frac{31}{35}$$

となります。

どうでしょうか？「少なくとも」という場合には、これを否定した事象を考えると簡潔に解決することが多いようです。

「少なくとも…」ときたら、否定した確率

Aの起こる確率=1-(Aの起こらない確率) …… 排反事象の定理

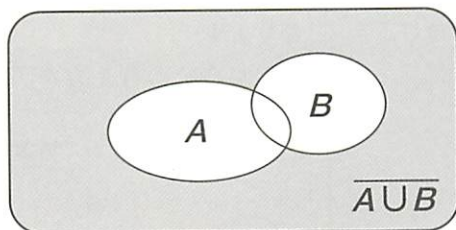


$$P(A) = 1 - P(\bar{A})$$

注) $\bar{A}$ は「Aでない」部分を表します。

参考

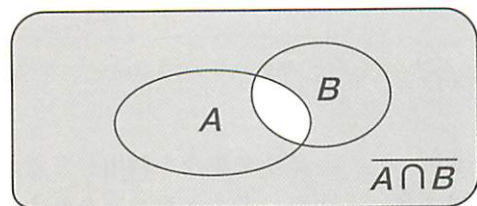
否定に関するド・モルガンの定理



$$\overline{A \cup B} = \bar{A} \cap \bar{B}$$

注) 読み方

「AまたはB」でない=
「Aでない」かつ「Bでない」



$$\overline{A \cap B} = \bar{A} \cup \bar{B}$$

注) 読み方

「AかつB」でない=
「Aでない」または「Bでない」

ここで、 $\bar{A}$ は集合Aの余事象、つまり、全体集合からAを除いた集合を表します。このド・モルガンの定理は、論理に関しても成立します。

ここが要点!

確率が最初に考えられたころは、明確に集合という考え方はしていなかったようです。しかし確率の考え方が深まり、また集合の考え方が数学全般に浸透するにつれて、確率の考えも集合を使って論じられるようになりました。

注) 集合の考え方を土台にして現代の数学は作られています。

やさしい解説

1つのサイコロを振って3の倍数の目(つまり、3と6)が出る確率はいくつですか。この質問に対して、答えは通常、次のようになります。

サイコロの目の出方は6通り、この中で3の倍数の目は3と6の2通り。よって、3の倍数の目が出る確率は、

$$\frac{2}{6} = \frac{1}{3}$$

となります。

この解答では、場合の数をカウントしただけですが、背景には次の集合の考え方があります。

$$U = \{1, 2, 3, 4, 5, 6\}$$

……出方のすべてからなる集合

$$A = \{3, 6\}$$

……確率を問題にしている集合

確率では、出方のすべてからなる集合を**標本空間**といい、標本空間の部分集合を**事象**といいます。

事象 A の起こる確率を $P(A)$ と書くことにすれば、先の解答は次のように

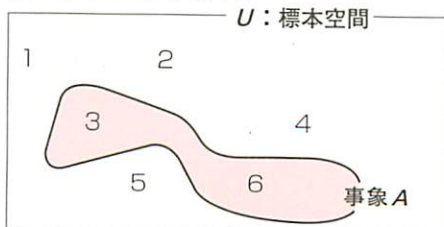
なります。つまり事象 A の起こる確率を $P(A)$ 、事象 A の要素の個数を $n(A)$ と書くことにすれば、

$$P(A) = \frac{n(A)}{n(U)} = \frac{A \text{ の要素の数}}{U \text{ の要素の数}}$$

$$\frac{2}{6} = \frac{1}{3}$$

となります。

この集合での考え方を図で書けば、次のようになります。



注) 集合のこのような図を「ベン図」といいます。

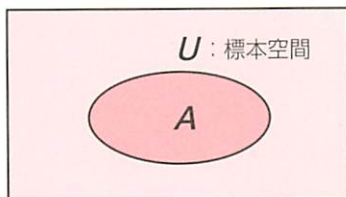
確率を集合を利用して考えることにより、考え方がスッキリするとともに集合のいろいろな性質が使えて便利になります。

確率の考えは集合を使ってまとめる

確率論

集合論

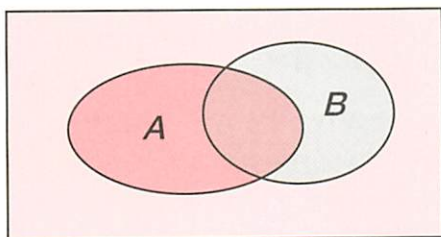
1 確率の定義(どれも同様な確からしさで起こる場合)



$$P(A) = \frac{n(A)}{n(U)}$$

$n(A)$ は、集合 A の要素の数。

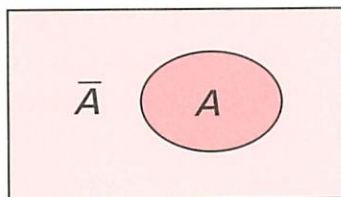
2 確率の加法定理 (§ 7)



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

事象 A または B のいずれかが起こる確率は、 A の起こる確率と B の起こる確率を加えて、 A と B がともに起こる確率を引けばよい。

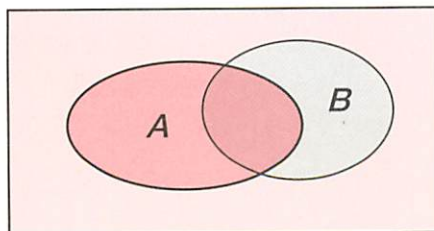
3 確率の排反事象の定理 (§ 11)



$$P(A) = 1 - P(\bar{A})$$

A の起こる確率は、1 から A の起こらない確率を引けばよい。

4 確率の乗法定理



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

事象 A と B がともに起こる確率は、 A の起こる確率 $P(A)$ に、 A が起きたときに B が起こる確率 $P_A(B)$ を掛け合わせればよい。

ここが要点!

1枚のコインを投げて、その表裏に着目する試行の場合は気にならなかったのですが、コインとサイコロを振る場合についてはちょっと気になることがあります。それは、この確率現象をきちんと数学に乗せて考えようとすると、「集合の掛け算」という考え方が必要になるということです。数学では集合の掛け算を「直積」といいます。

やさしい解説

集合とは、ある条件を満たすものの集まりです。この集まり同士の掛け算とはどういうことなのでしょう。このことについて調べてみましょう。

たとえば、2つの数の集合を、

$$A = \{1, 2, 3\}, B = \{4, 5\}$$

とします。このとき、これら2つの集合を掛け合わせたもの $A \times B$ は次の集合であるとし、これを直積と呼びます。

$$A \times B = \left\{ \begin{array}{cc} (1, 4) & (1, 5) \\ (2, 4) & (2, 5) \\ (3, 4) & (3, 5) \end{array} \right\}$$

つまり、「**Aの要素を左、Bの要素を右に配置した順序のついた組のすべての集まりを集合AとBの直積**」というのです。

なお、普通の数の掛け算では掛ける数と掛けられる数を入れ替えても結果は同じでした(交換法則が成り立つといえます)。ところが、直積には交換

すると、違った集合Aになってしまいます。つまり、

$$B \times A = \left\{ \begin{array}{ccc} (4, 1) & (4, 2) & (4, 3) \\ (5, 1) & (5, 2) & (5, 3) \end{array} \right\}$$

となりますから、この場合、

$$A \times B \neq B \times A$$

です。

それでは、なぜ「直積」にこだわるのでしょうか。それはこれを使うことによって、標本空間が簡潔に表現されるからです。たとえば、コインを投げてサイコロを振る試行においては、標本空間は $A \times B$ と書けます。ただし、

$$A = \{\text{表}, \text{裏}\}, B = \{1, 2, 3, 4, 5, 6\}$$

です。

もし、サイコロを振ってコインを投げる試行であれば標本空間は $B \times A$ と書けます。また、もしサイコロを100回振る試行であればこの標本空間は、

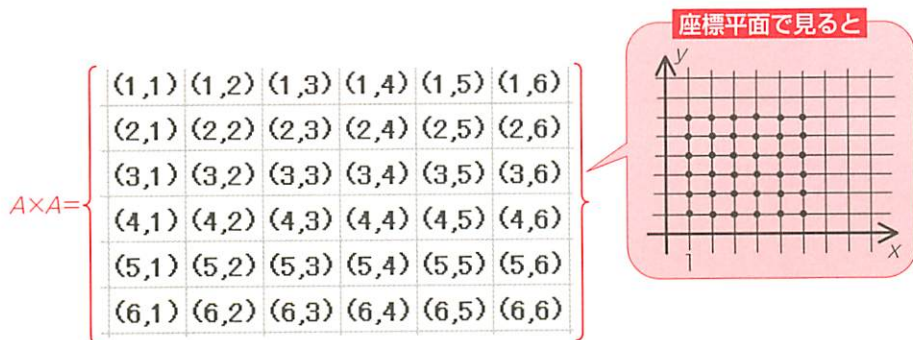
$$A \times A \times A \times \cdots \times A \times A$$

というようにAの100個の積で書けます。

直積は標本空間の表現に便利

1 サイコロを2回振る試行の標本空間は $A \times A$

ただし、 $A = \{1, 2, 3, 4, 5, 6\}$



2 サイコロを振り、コインを投げる試行の標本空間は $A \times B$ コインを投げてサイコロを振る試行の標本空間は $B \times A$

ただし、 $A = \{1, 2, 3, 4, 5, 6\}$,

$B = \{\text{表}, \text{裏}\}$

$A \times B =$

(1,表)	(1,裏)
(2,表)	(2,裏)
(3,表)	(3,裏)
(4,表)	(4,裏)
(5,表)	(5,裏)
(6,表)	(6,裏)

$B \times A =$

(表,1)	(表,2)	(表,3)	(表,4)	(表,5)	(表,6)
(裏,1)	(裏,2)	(裏,3)	(裏,4)	(裏,5)	(裏,6)

3 直積 $A \times B$ の要素の個数は、おのこの要素の個数の積。 $n(A \times B) = n(B \times A) = n(A) \times n(B)$

ただし、 $n(A)$ は集合 A の要素の数

ここが要点！

確率では「一方が起これば他方が起こらない」とき、2つの事象は排反であるといえます。つまり、「2つが同時に起こらない」ということです。この結果、AかBのいずれか一方の起こる確率は、単純におのの起こる確率の和になります。

やさしい解説

2つの事象AとBが排反であるとは、一方が起これば他方は起こらないということです。たとえば、コインを投げて表が出れば裏は出ないということです。つまり、「同時には起こり得ないということです。集合で表現すれば、事象AとBに共通部分がない」ということです。つまり、

$$A \cap B = \phi$$

です。ここで、 ϕ (ファイ) は空集合といい、要素を1つも持たない集合です。

3つの事象AとBとCが排反であることも同様です。つまり、一方が起これば他方のいずれも起こらないということです。

事象AとBのいずれかが起こる確率を求めるには、一般には次の確率の加法定理を使います。

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

ところが、事象AとBが排反であれば、

$$P(A \cap B) = P(\phi) = 0$$

なので、計算が次のように簡単になります。

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

この排反という条件は、事象が3個以上あるときには、すごくありがたいのです。事象がA、B、Cの3個あるときは、事象AとBとCのいずれかが起こる確率を求めるには、次の複雑な式を使います。

$$\begin{aligned} P(A \cup B \cup C) &= P(A) + P(B) + P(C) \\ &\quad - P(A \cap B) - P(B \cap C) - P(C \cap A) \\ &\quad + P(A \cap B \cap C) \end{aligned}$$

ところが、AとBとCがお互いに排反であれば、

$P(A \cup B \cup C) = P(A) + P(B) + P(C)$ と非常に簡単になります。このように、排反は確率計算をする上では非常にありがたい条件なのです。

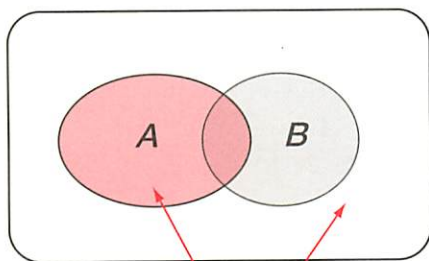


排反は確率計算が楽になるすばらしい条件

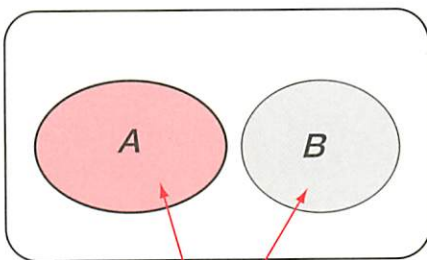
1 2つの事象が排反とは2つが分離

$$A \cap B = \phi$$

あちら立てれば、
こちらが立たず



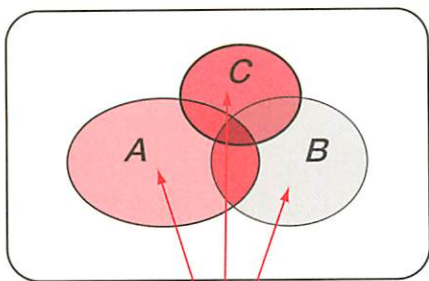
排反でない



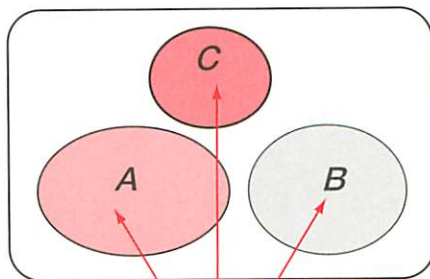
排反

2 3つの事象が排反とは3つが互いに分離

$$A \cap B = \phi, B \cap C = \phi, C \cap A = \phi$$



排反でない



排反

ここが要点!

確率の勉強をしていると、しばしば「影響する」とかしないとかいう言葉に出くわします。しかし、数学の立場からすると、一方が他方に影響しているかどうかは証明できません。数学では「独立」ということを式で定義して、この式にあてはまるかどうかで判定します。

やさしい解説

2枚のコイン α と β を投げて、表裏の出方に着目してみましょう。多くの人は、一方のコインの表、裏の出方は他方の表、裏の出方に影響しないと思っています。しかし、このことは数学では証明できません。そこで数学では次のようにして**独立**ということを経定義して、影響論の考えを取り込んでいます。つまり、

「 $P_A(B) = P(B)$ ……①が成立するとき、2つの事象は独立である」

とするのです。ここで、 $P_A(B)$ とは、事象 A が起きたときに事象 B が起こる確率という意味で、式で書けば次のようになります。

$$P_A(B) = \frac{P(A \cap B)}{P(A)} \dots\dots\dots ②$$

これは、事象 A をあたかも標本空間と考えたときの事象 B の起こる確率ということです。

以上の準備のもとに、先程の2つのコインの話にもどりましょう。

いま、コイン α が表であるという事象を A 、コイン β が表であるという事象を B とします。

A		
(α 表, β 表)	(α 裏, β 表)	B
(α 表, β 裏)	(α 裏, β 裏)	

2枚のコインを投げれば、上の4通りは表裏の出方は影響しないので等確率 ($1/4$) だと考えられます。このとき、式①が成立し、 A と B は独立と判断されます。しかし、この4通りが次の確率で出た場合は式①は成立しません。

$$(\alpha \text{ 表}, \beta \text{ 表}) \frac{1}{6}, (\alpha \text{ 裏}, \beta \text{ 表}) \frac{2}{6}$$

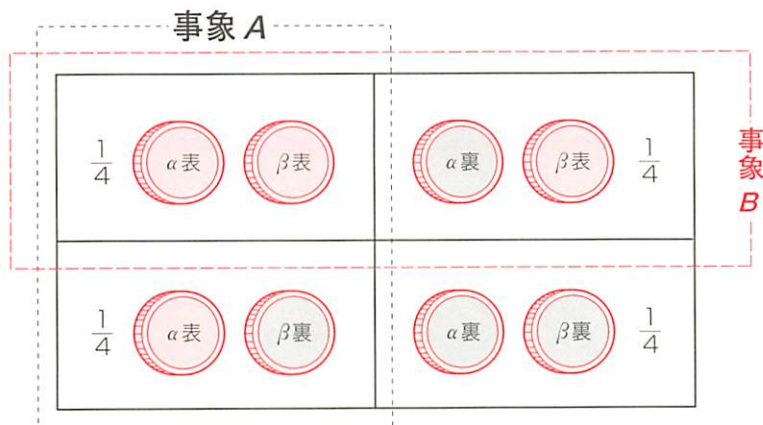
$$(\alpha \text{ 表}, \beta \text{ 裏}) \frac{2}{6}, (\alpha \text{ 裏}, \beta \text{ 裏}) \frac{1}{6}$$

したがって、数学では A と B は独立でないと判定します。 A と B が影響しあって2枚とも表や、2枚とも裏が出にくくなったからだと思われます。

このように、式①による独立の定義は、感覚的な独立と結びついたすてきな定義だと思われます。

数学の独立は感覚的な影響論とマッチする

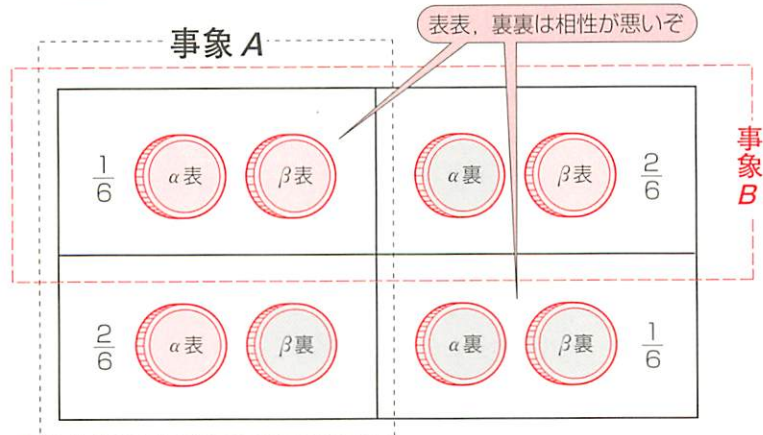
1 事象 A と B は独立



なぜか

$$P(B) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}, P_A(B) = \frac{P(A \cap B)}{P(A)} = \frac{\frac{1}{4}}{\frac{1}{4} + \frac{1}{4}} = \frac{1}{2}. \text{ よって } P(B) = P_A(B)$$

2 事象 A と B は独立でない



なぜか

$$P(B) = \frac{1}{6} + \frac{2}{6} = \frac{1}{2}, P_A(B) = \frac{P(A \cap B)}{P(A)} = \frac{\frac{1}{6}}{\frac{1}{6} + \frac{2}{6}} = \frac{1}{3}. \text{ よって } P(B) \neq P_A(B)$$

ここが要点!

確率では「一方が起これば他方が起こらない」とき、2つの事象は排反であるといえます。また、「一方の起こる起こらないが他方に影響しない」とき、2つの事象は独立であるといえます。この2つは、よく混同されるので注意してください。

やさしい解説

排反と独立は、確率で非常によく使われる考え方です。ところが、この双方を取り違えて解釈している人も少なくありません。

2つの事象 A と B が排反であるとは、「一方が起これば他方は起こらない」ということです。つまり、同時に起こり得ないということです。集合でいえば、事象 A と B に共通部分がないということです。つまり、 $A \cap B = \phi$ です。

この結果、2つの事象 A 、 B が排反であれば、 A と B の少なくとも一方が起こる確率は、おのこの確率を足せばよいということになります。

これに対して独立とは、一方の起こる起こらないが他方に影響しないということです。しかし、この影響するかわからないかは確かめることが困難です。そこで、事象 A が事象 B に影響しないということを次式で表現したのです。

$$P_A(B) = P(B) \cdots \cdots \textcircled{1}$$

ここで、 $P_A(B)$ は事象 A が起きたときに事象 B が起こる確率です。この式が成立しているときには、

$$P_{\bar{A}}(B) = P(B) \cdots \cdots \textcircled{2}$$

が成立します。したがって式①②から、

$$P_A(B) = P_{\bar{A}}(B)$$

となります。 A が起きても起きなくても B の起こる確率は等しいというわけですから、式①の成立をもって影響しないと考えようというわけです。

この式①から、次式③が成立することが導き出せます。

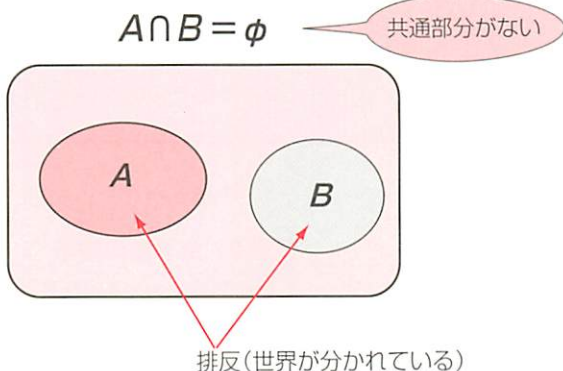
$$P(A \cap B) = P(A) P(B) \cdots \cdots \textcircled{3}$$

したがって、この式③をもって2つの事象 A 、 B が独立だということもあります。

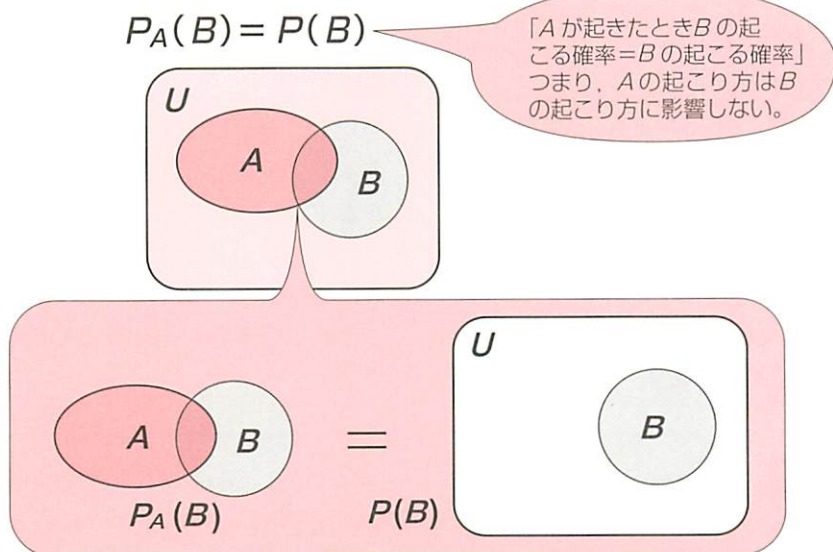
式③によると、2つの事象 A 、 B が独立であれば、 A と B がともに起こる確率は、おのこの確率を掛ければよいということになります。独立ということと掛け合わせることは密接な関係があります。

排反と独立は別なもの

1 排反とは



2 独立とは



注1) $P_A(B)$ とは、Aが起きたときBの起こる確率という意味で、Aをあたかも標本空間と考えたときのBの起こる確率のこと。つまり、

$$P_A(B) = \frac{P(A \cap B)}{P(A)} \quad \text{なお、} P_A(B) \text{は } P(B|A) \text{とも書きます。}$$

2) 基本的には $P_A(B) = P(B)$ は、 $P(A \cap B) = P(A)P(B)$ と同じと見なせます。

ここが要点!

ユークリッド幾何学は、いくつかの公理を決めて、そこから純粋に演繹的考察で一連の正しい図形の性質(定理)を導いています。不確かな事柄を考察とする確率をユークリッド幾何学のように公理を決めて攻めたのがロシアのコルモゴロフ(1903~1989)です。

やさしい解説

確率の公理的定義を紹介すると、次のようになります。

標本空間(全体集合) S が与えられているとき、各事象(集合) A, B に対して次の条件(公理)を満たす数 $P(A)$ を対応させます。

$$(1) P(A) \geq 0$$

$$(2) P(S) = 1$$

$$(3) A \cap B = \phi \text{ のとき}$$

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

このとき、 $P(A)$ を事象 A の確率といいます。

以上の3個の公理のみから、次のいろいろな定理が導き出されます。

定理1 空事象の確率

$$P(\phi) = 0$$

定理2 余事象の確率

$$P(\bar{A}) = 1 - P(A)$$

定理3

$$0 \leq P(A) \leq 1$$

定理4

$$A \subset B \text{ ならば } P(A) \leq P(B)$$

定理5 確率の加法定理

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

定理6

以下、確率に関するいろいろな定理をイモヅル式にどんどん導くことができます。

(1), (2), (3) では、確率らしきことには触れていません。それなのに、ここから確率論が作られるのは驚きです。

参考

場合の数による確率

場合の数による確率は、公理的確率の特別な場合である。つまり、 $S = \{e_1, e_2, e_3, \dots, e_N\}$ のとき、各根元事象に等確率 $1/N$ を与えたものです。つまり、

$$P(\{e_1\}) = P(\{e_2\}) = \dots = P(\{e_N\}) = \frac{1}{N} \quad \text{このとき、たとえば } A = \{e_1, e_2, e_3\} \text{ とすると、}$$

$$P(A) = P(\{e_1\}) + P(\{e_2\}) + P(\{e_3\}) = \frac{1}{N} + \frac{1}{N} + \frac{1}{N} = \frac{3}{N} = \frac{\text{要素} A \text{ の要素の数}}{\text{標本空間の要素の数}}$$

公理的確率論 ～ ユークリッド幾何学

確率の公理

- (1) $P(A) \geq 0$
- (2) $P(S) = 1$
- (3) $A \cap B = \phi$ のとき
 $P(A \cup B) = P(A) + P(B)$

演繹

余事象の確率

$$P(\bar{A}) = 1 - P(A)$$

確率は0以上1以下

$$0 \leq P(A) \leq 1$$

 $A \subset B$ ならば

$$P(A) \leq P(B)$$

.....
.....

確率の加法定理

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

.....
.....

もはや、「場合の数」
や「同程度」などの
言葉はありません。

ユークリッド幾何学の公理

- (1) 2点 A, B を通る直線はただ 1 つある。
- (2) 有限な直線(線分)を直線に延長することができる。
- (3) 任意の中心と半径を持つ円が描ける。
- (4)
- (5)

演繹

二等辺三角形
の底角は等しい

3つの辺がそれぞれ等しい三角形は合同



ピタゴラスの定理

.....
.....
..........
.....

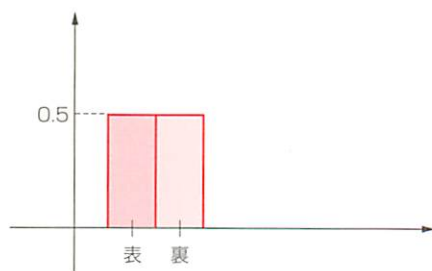
注) 演繹(えんえき)とは、一般的なものから特殊なものへの推論。三段論法など。

ここが要点！

コインを投げて表とか裏とか，トランプからカードを抜きとって，ハートとかスペードとかが出る確率を考えると，このままでは不十分なことがあります。たとえば，それぞれの出やすさをグラフ化するとき，表とか裏とかハートとかスペードとかを座標軸上にどうとてよいか困ってしまいます。そこで，確率が与えられた変数，つまり，確率変数というものを考えることにしました。

やさしい解説

1枚のコインを投げた場合，通常，表が出る確率と裏の出る確率はともに等しくて $1/2$ と考えられます。このことをグラフ化すると，次のようになります。



しかし，横座標が「表」や「裏」では計算をする上で何かと不都合です。

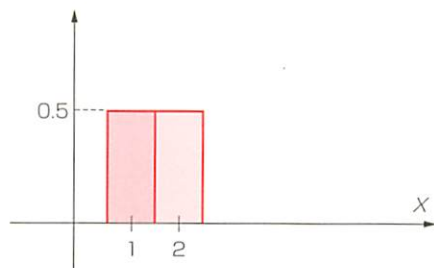
そこで，ある試行に対して，その値が決まる**偶然変量 X** を導入することになります。たとえば，1枚のコインを投げる試行に対して表が出たら1，裏が出たら2という値をとる偶然変量を X としてみます。すると， $X=1$ である確率は0.5， $X=2$ である確率も0.5と

なります。このことを

$$P(X=1) = 0.5, P(X=2) = 0.5$$

などと書きます。このとき， X は複数の値をとる数，つまり変数です。しかも，単なる変数ではありません。確率が与えられた変数なのです。

数学では，このように確率が付与された変数を**確率変数**と呼びます。確率変数 X の値を横座標に，そのときの確率を縦座標にとることにより，確率変数 X がどんな確率をとるかを座標平面上にグラフで表現できます。



このように確率変数に確率が分配されたようすを表すグラフを**確率分布のグラフ**といいます。

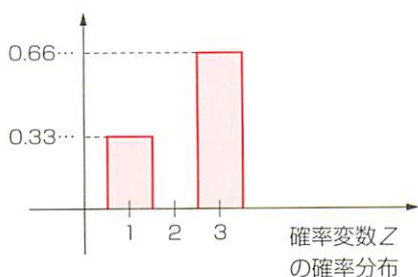
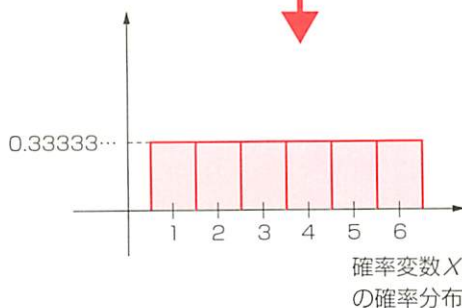
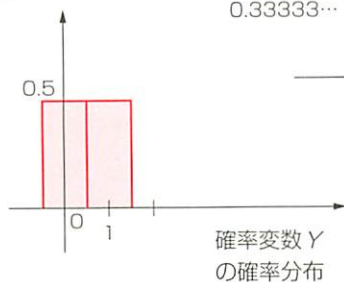
確率現象を数値化して把握するのが確率変数

1 サイコロを振って目の出方に着目する試行でも、いろいろな確率変数が考えられる

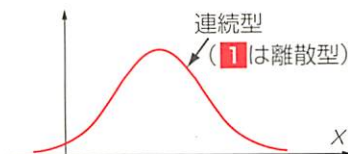
出た目の数を X

出た目が2の倍数なら
 $Y = 1$, そうでなければ $Y = 0$

出た目が3の倍数なら
 $Z = 1$, そうでなければ $Z = 0$



2 確率変数には離散型と連続型がある



ここが要点!

「1回の試行で事象 A の起こる確率が p であるとする。この試行を独立に n 回繰り返すとき、事象 A が r 回起こる確率は、

$${}_nC_rp^rq^{n-r} \quad (\text{ただし、} q=1-p)$$
 となる」これを独立試行の定理といいます。何回も繰り返される確率現象の多くは、この定理で説明できます。

やさしい解説

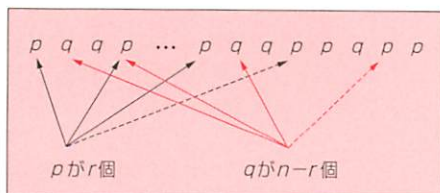
コインを10回投げて6回表が出る確率はどのくらいでしょうか。試験問題1問を解ける確率が0.6であるとき、5問中3問以上解ける確率はどのくらいでしょうか。麻雀で勝つ確率が0.55の人が100回賭をしたときの儲けはどのくらいになるのでしょうか。……

これらは、いずれも独立試行の定理を用いて求めることができます。

1回の試行で事象 A の起こる確率が p であるとき、この試行を n 回行い、そのうち事象 A が r 回起きたとすれば、その確率は、

$$p^r q^{n-r}$$

となります。ただし、 $q=1-p$



ここで、 n 回中、事象 A が r 回起こ

る起こり方は全部で、 ${}_nC_r$ 通りあります。

これは n 個の座席があって、 r 個の座席で事象 A が起こったと考えると、座席の選び方は ${}_nC_r$ 通りであることからわかります。これら ${}_nC_r$ 通りについては、どれか一通りが起これば他は起こらない(排反)ので、求める確率は、

$${}_nC_rp^rq^{n-r} \dots \dots \dots \textcircled{1}$$

ということになります。これを**独立試行の定理**といいます。

(例) 表が出る確率が0.6であるコインを10回投げたとき、表が r 回出る確率は、独立試行の定理より、

$${}_{10}C_r 0.6^r 0.4^{10-r}$$
 となります。

なお、一般に確率変数 X が r をとる確率 $P(X=r)$ が式①のとき、この確率変数の分布が有名な二項分布です。

繰り返し現象には独立試行の定理

1 n 回中 r 回◎ (1 回で起こる確率は p) が起こる確率は ${}_nC_rp^rq^{n-r}$

「 n 個の席から
 r 個選んで、そ
こを◎とする」
と考える。

${}_nC_r$
通りの出方がある

1 2 3	i	n 回	
○○○○×○○		×××○○×	この確率は p^rq^{n-r}
○○○○×××		○××××	この確率は p^rq^{n-r}
○×○○×○○		×○×××	この確率は p^rq^{n-r}
○○○×○○○		××○○×	この確率は p^rq^{n-r}
○○×○○○○		×××○○×	この確率は p^rq^{n-r}
.....			
.....			
.....			
××××××		×○○○×	この確率は p^rq^{n-r}
××○×××		○×××○	この確率は p^rq^{n-r}
××○○○×		○○○○×	この確率は p^rq^{n-r}
××××××		○○××○	この確率は p^rq^{n-r}
××××××		○○○○○	この確率は p^rq^{n-r}

合計 ${}_nC_rp^rq^{n-r}$

2 繰り返し現象の確率は ${}_nC_rp^rq^{n-r}$ を使う

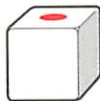
1 回目

2 回目

3 回目

.....

n 回目



.....

n 回中、 r 回 1 が出る確率！



.....

n 回中、 r 回表が出る確率！



.....

n 回中、 r 回成功する確率！

ここが要点！

確率論は、17世紀に賭に関するパスカルとフェルマーの議論から始まり、19世紀にラプラスにより「同様な確からしさ」を前提とする数学的確率が作られました。しかし、この前提の正しさを示すことは難しいので、その後、経験にもとづく統計的確率が作られました。

注) 数学的確率と統計的確率の難点を克服するものとして公理論的確率が作られました。

やさしい解説

(1) 数学的確率 (先験的確率, 古典的確率)

ある試行によって起こることが可能
なすべての結果 (根元事象) の総数が
 N であり、どの場合も同様な確から
しさで起こるものとする。このうち、あ
る事象 (事柄) A の起こる場合の総数
が $n(A)$ ならば、その事象 A の起こる
確率は、

$$P(A) = \frac{\text{事象 } A \text{ の場合の数}}{\text{全体の場合の数}} = \frac{n(A)}{N}$$

であるとしています。

このように考える確率論を数学的確率または**先験的確率**、**古典的確率**といいます。これは、順列・組合せ論を基礎とする確率論です。

(例) 赤3個、白2個合わせて5個の玉から2個取り出すとき、赤1個、白1個を取り出す数学的確率は、

$$\frac{{}_3C_1 \times {}_2C_1}{{}_5C_2} = \frac{3 \times 2}{10} = \frac{3}{5}$$

(2) 統計的確率 (経験的確率)

数学的確率の難点は、「根元事象がどれも同様に確からしく起こる」ことを確かめる合理的方法を見出すことが難しいということです。この難点を克服したのが、次の統計的確率です。

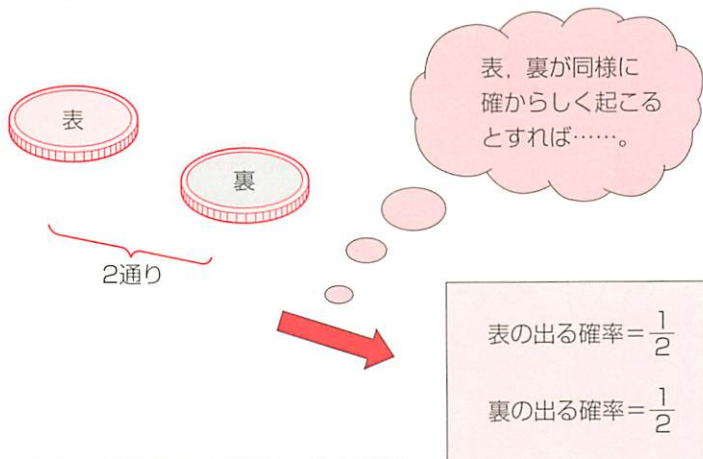
これは、現実の事象を観察して確率を推定しようというものです。つまり、ある試行を n 回繰り返したとき、事象 A が a 回起こったとします。試行回数 n を大きくしたとき、相対度数 a/n の値がどんどん一定の値に近づけば、その値を**統計的確率**または**経験的確率**といいます。しかし、この考え方も無限に試行を続けられないという難点があります。

(例) 画鋲を10万回投げたら51000回針が上を向いた。よって、画鋲が上を向く統計的確率は、

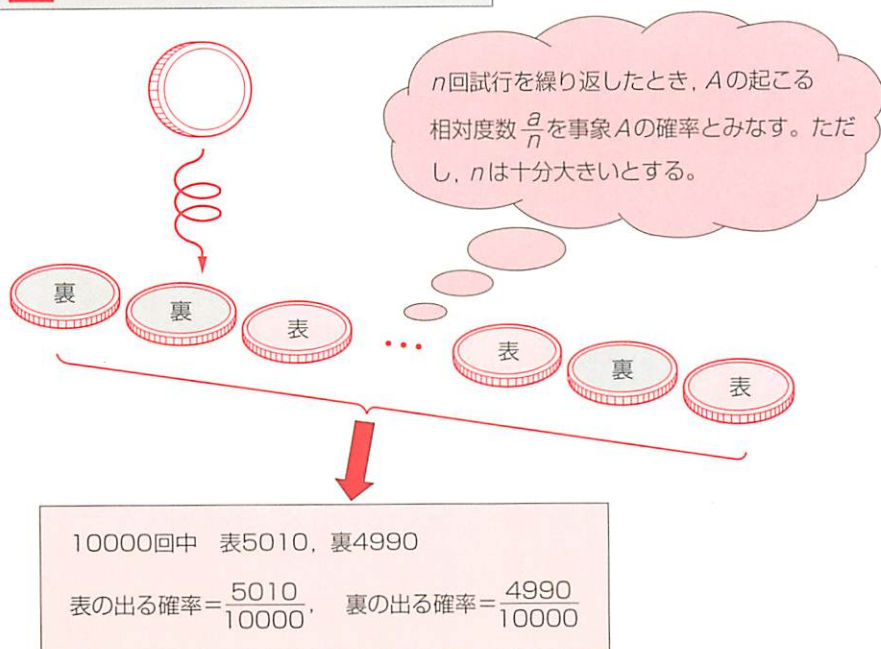
$$\frac{51000}{100000} = \frac{51}{100} = 0.51$$

経験しなくてもわかる数学的確率， 豊富な経験でわかる統計的確率

1 数学的確率



2 統計的確率



ここが要点！

10円硬貨を何回か投げると、そのたびごとに表の出る相対度数は大きく変化します。これは多くの人が経験しているところです。ところが、たくさん投げ続けると、表の出る相対度数はほぼ0.5に落ち着きます。この性質を「大数の法則」といいます。

やさしい解説

ある事柄が一見、偶然に起きたりデタラメに起きたりしているように見えても、観察回数を増やしていくと、その事柄の起こる相対度数は一定の値に近づくことがわかります。

確率の言葉でいい換えると、次のようになります。

「ある事象 A の起こる確率を p 、 n 回の独立試行で A が起きた相対度数を r/n とするとき、 n が十分大きくなれば、 r/n がほぼ p に等しいことはほとんど確実である」

これが**大数の法則**と呼ばれるものです。

この法則によれば、試行回数をどんどん増やしていくと、ある事象 A の起こる相対度数は事象 A の起こる固有の確率 p に分け隔てなく近づいていくということです。これは10円硬貨の場合だけではなく、サイコロを振る場合でも同じです。1の目の出た相

対度数に着目すると、最初はバラツキがありますが、振り続けるとその相対度数は $1/6$ に近づいていくのです。

この大数の法則からは、どんな場合でも逃れることはできません。「俺は賭事にはついてる」というギャンブラーも同じことです。

なお、この「大数の法則」によって、数学的確率と統計的確率が結びつきます。

注) 一般に「大数の法則」と呼ばれるものは、先の法則を一般化した次の法則です。

「互いに独立で同一の確率分布（平均は m ）を持つ確率変数の列

$$X_1, X_2, \dots, X_n$$

において、 $n \rightarrow \infty$ のとき、

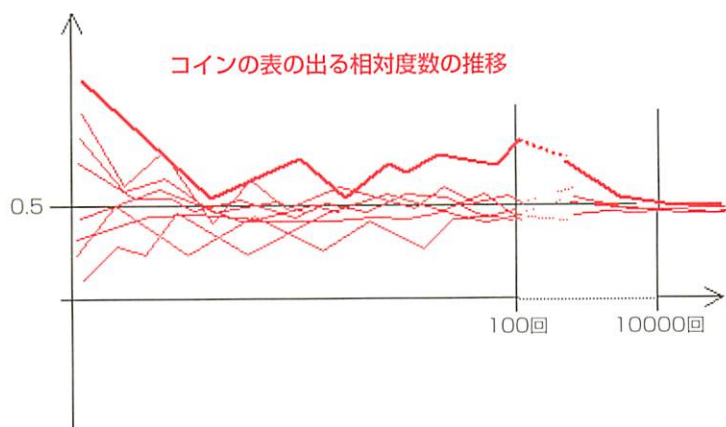
$$\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n}$$

は、 m に限りなく近づく」

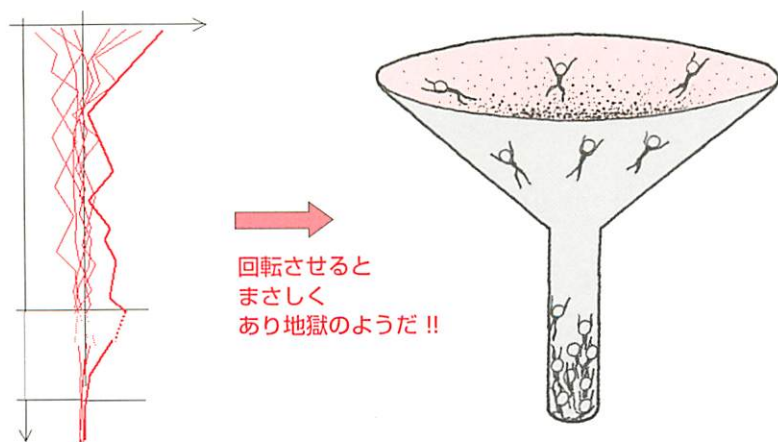
ちょっと難しいですね。

「大数の法則」には逆らえない

1 「大数の法則」をグラフで実感



2 回転してみると、まさしくあり地獄



§ 19 独立にもいろいろ

ここが要点!

「独立」をいい表した確率の用語に「事象の独立」、「試行の独立」、「確率変数の独立」の3つがあります。いずれも「影響しない」ということをいい表したのですが、中身の違いに注意しましょう。

やさしい解説

(1) 事象の独立

「2つの事象があり、一方の起こる確率が他方の起こる確率に影響しないとき、これらの事象は独立だ」といいます。たとえば、甲、乙2個のサイコロを投げたとき、甲が2以下の目が出るという事象をA、乙が6の目が出るという事象をBとすると、2つの事象A、Bは独立です。

注) 影響するかどうかは、数学では判定できません。そこで、数学では事象の独立を式で定義します。このことについては§12参照。

(2) 試行の独立

「一方の試行の結果が他方の試行の結果に影響しないとき、2つの試行は独立である」といいます。たとえば、サイコロを振って出た目の数に着目する試行 α と、コインを投げてその表裏に着目する試行 β は独立です。

注) 式の上では、2つの試行を組み合わせた試行の確率は、おのおのの試行の確率の積に等しいということです。たとえば、コインとサイ

コロを振る試行で、コインが表でサイコロが1になる確率は、一方の出方が他方に影響しないので独立だから、おのおのの確率を掛け合わせた

$$\frac{1}{2} \times \frac{1}{6} = \frac{1}{12}$$

となります。

(3) 確率変数の独立

「同一の標本空間で定義された2つの確率変数があるとき、一方の確率変数のとる値が他方の確率変数のとる値に影響しなければ、これらの確率変数は独立だ」といいます。たとえば、コインを2回投げる試行で*i*回目に出れば1、裏が出れば0という確率変数 X_i とすると2つの確率変数 X_1, X_2 は独立です。

注) 式の上では、次のようになります。

確率変数 X と Y があるとき、 X が a で Y が b という値をとる確率 $P(X=a, Y=b)$ は、 X が a をとる確率 $P(X=a)$ と Y が b をとる確率 $P(Y=b)$ を掛け合わせたものに等しいということです。つまり、

$$P(X=a, Y=b) = P(X=a) \times P(Y=b)$$

これが X, Y のとる任意の a, b について成り立つとき X と Y は独立なのです。

独立はどれも掛け算と仲がいい

1 事象の独立(甲, 乙2つのサイコロを振った場合)

直観!!

事象AとB
は影響しないから独立

乙 \ 甲	1	2	3	4	5	6
1	(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
2	(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
3	(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
4	(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
5	(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
6	(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

事象A: 甲が2以下

事象B: 乙が1

数式!!

$$P(A) = \frac{12}{36}, P(B) = \frac{6}{36}$$

$$P(A \cap B) = \frac{2}{36} \text{ となり,}$$

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

が成立する。よって独立

2 試行の独立(サイコロ甲を振る試行α, サイコロ乙を振る試行β)

直観!!

試行αとβは
影響しないから独立

乙 \ 甲	1	2	3	4	5	6
1	(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
2	(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
3	(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
4	(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
5	(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
6	(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

数式!!

たとえば, 甲が3, 乙が2である確率は $\frac{1}{36}$

甲が3である確率と, 乙が2である確率を掛けて,

$$\frac{1}{6} \times \frac{1}{6} = \frac{1}{36}$$

この2つは等しいからαとβは独立

3 確率変数の独立(サイコロ甲の目をX, サイコロ乙の目をY)

直観!!

X, Yの値は
お互いに影響しないから独立

		X=1	X=2	X=3	X=4	X=5	X=6
Y=1	乙 \ 甲	1	2	3	4	5	6
Y=2	1	(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
Y=3	2	(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
Y=4	3	(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
Y=5	4	(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
Y=6	5	(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
6	6	(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

数式!!

たとえば, Xが3, Yが2である確率は $\frac{1}{36}$

ここで, Xが3の確率

$$\frac{6}{36} = \frac{1}{6} \text{ と, Yが3の確率}$$

$$\frac{6}{36} = \frac{1}{6} \text{ を掛けると } \frac{1}{36} \text{ となり,}$$

$$P(X=3, Y=2) =$$

$$P(X=3) \times P(Y=2)$$

となり, XとYは独立

ここが要点!

同じ1万分の1の確率でも、それが個人にとってすてきな確率現象とそうでない場合とでは、気持ちの上での起きやすさの度合いは大分違うようです。冷静な判断をしたいものです。

やさしい解説

「癌の再発は100分の1の確率です」といわれると、生きた心地がしないものです。再発は十分にあり得ると思ひ、いろいろな準備にとりかかります。

ところが、「100分の1の確率で1万円が当たる」場合には、「ひょっとしたら当たるかもしれない」と思ひます。けっして、当たる可能性がかなり高いとは思わないうです。このように同じ確率でも、それが起こることが個人にとっていいことか悪いことかで、起こりやすさの度合いが気分の上で大分違ってくるようです。

また、数字のトリックで確率がぼやけてしまうこともあります。たとえば、次の宝くじ(イ)、(ロ)のどちらを買うかと聞かれると、多くの人は(ロ)のくじを購入します。理由は、いままで(イ)のような番号のくじが当たった試しがないからと答えます。

11組 111111 ……………(イ)

38組 152874 ……………(ロ)

しかし、確率の考えからすれば、宝くじに八百長でもない限り、(イ)のくじも(ロ)のくじも当たる確率は変わらないはずです。したがって、どちらを買ってもいいことになります。でも、(イ)の方が心理的には当たりにくいと思われてしまいます。(イ)が当たらないのであれば(ロ)も当たれませんから、宝くじは買わない方がいいのですが……。

確率現象で何かうまい話に出くわしたら、それと等確率の極端な事例を思ひ浮かべることが大事なかもしれません。

また、スポーツの試合で陣地をくじで決めるのに、そのくじをどちらが先に引くかをくじで決めるということがあります。

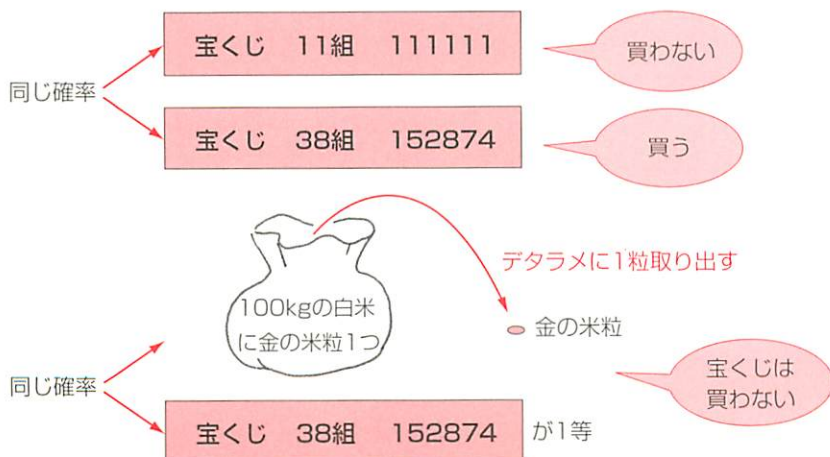
当たりくじを引く確率は、くじを引く順番に関係ない(§ 34 参照)ことが理論的にはわかっていますが、このように確率現象を幾重にも組み合わせる使うことが珍しくありません。これも多分に心理的な問題かもしれません。

同じ確率でも気分で左右される

- 1 同じ確率でも、嫌なことは確率が大きく感じられる。
いいこと、そうあって欲しいことは確率が少なく感じられる。



- 2 心理的確率を逆利用して判断を再確認



ここが要点！

フランスの貴族シュバリエ・ド・メレが「人間は考える葦である」という名言をはいたパスカルに次の質問をしました。「A、Bの2人が1万円ずつ賭金を出し合ってゲームをし、最初に3勝したほうが賭金の2万円をもらうことにしました。ところがAが2勝、Bが1勝したところで、このゲームを中止しなければならなくなりました。このとき、賭金2万円をどう分配したら合理的でしょうか。ただし、2人のゲームの技量は同等とします」と。これが、確率論のルーツだといわれています。

やさしい解説

このメレの質問に対して、パスカルは次の解答を用意しました。

3回勝負がついたところで(Aが2勝1敗)、ゲームが中断したわけです。さらにゲームを続行すると仮定すると、次の2つの場合があります。

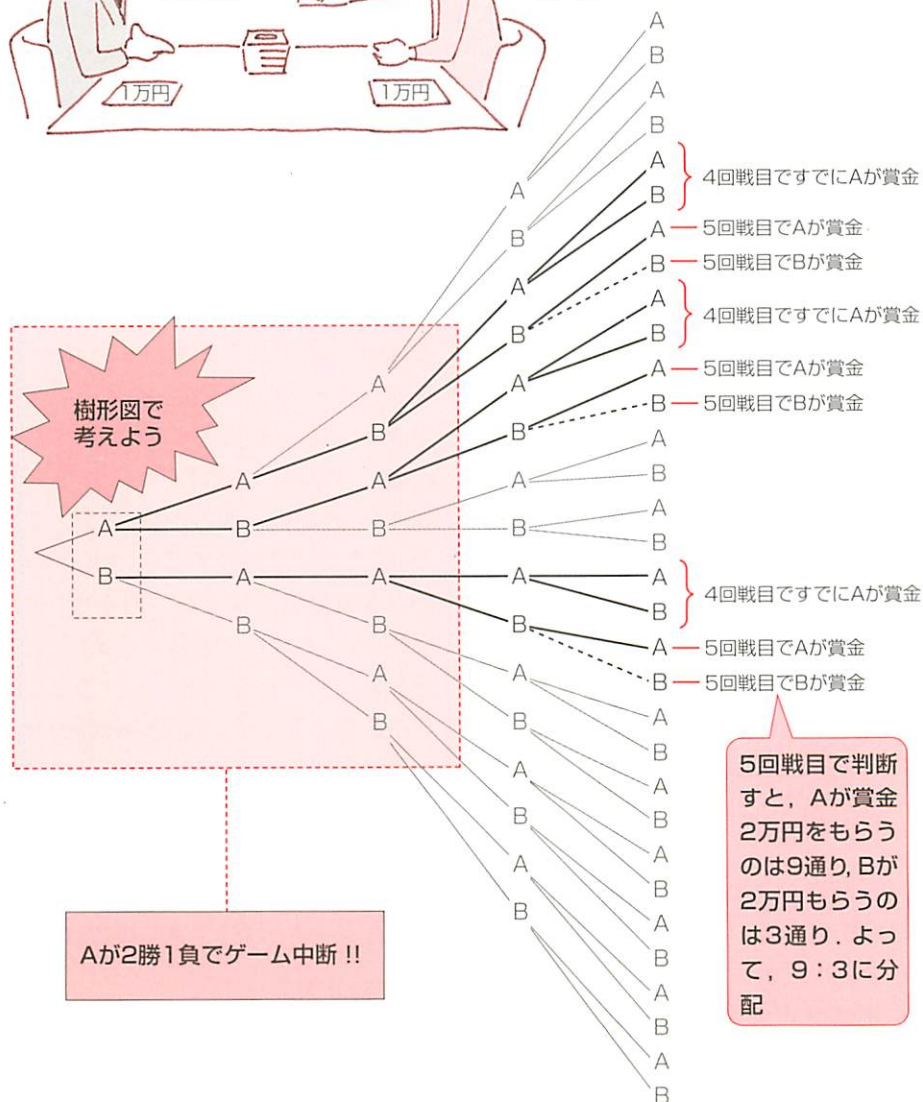
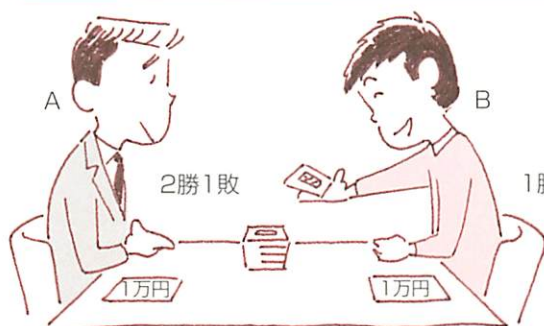
- (1) 4回目の勝負にAが勝てば、AはBより先に3勝することになります。したがって、賭金の2万円は全額Aのものになります。
- (2) 4回目の勝負にAが負ければA、Bともに2勝2敗の同成績になります。したがって、もしここでゲームを打ち切るとすれば、賭金2万円を折半して1万円ずつ受けとることになります。したがって、ゲームを中止した時点

で賭金を分配しようとする場合には、A、Bに次のように分けることにします。

「Aは次の勝負に勝っても負けても1万円はもらえます。残りの1万円は、次の勝負にAが勝てばAのもの、Bが勝てばBのものです。ところで、次の勝負にAが勝つかBが勝つかは五分五分です。したがって、まずAにAのものであることが確実な1万円をあげます。そして残りの1万円を5千円ずつA、Bに分けます。こうしてAに15000円、Bには5000円を分配するのが、この場合、最も妥当な賭金分配方法である……」

以上がパスカルの解答の要約です。いまならば、中学生や高校生でもこのような解答にたどり着きますが、当時としては大変な問題でした。

いまでは御法度の賭事が確率論の始まりとは!!



ここが要点！

コインを2枚投げるとき、表裏の出方は、次の3通りです。「2枚とも表」、「1枚が表、もう1枚が裏」、「2枚とも裏」。このことをもとにフランスの著名な数学者ダランベールは「少なくとも1枚が表」である確率を $2/3$ と判断しました。ところが、実際にコインを何回も振って調べてみると違う値になります。したがって、これはダランベールの判断ミスでした。

やさしい解説

実際に甲、乙2枚のコインを1000回投げてみました。その結果ですが、次のようになりました。

甲 \ 乙	表	裏
表	232	247
裏	266	255

相対度数を調べてみると、次のようになります。

甲 \ 乙	表	裏
表	0.232	0.247
裏	0.266	0.255

試行回数が1000回で、ある程度多いので相対度数がほぼ実際の確率に近いものと思われます。よって、少なくとも1枚が表である確率は、{甲表、乙表}、{甲表、乙裏}、{甲裏、乙表}の3つの事象の確率を足して、

$$0.232 + 0.247 + 0.266 = 0.734$$

となります。この値は、

$$\frac{2}{3} = 0.6666 \dots$$

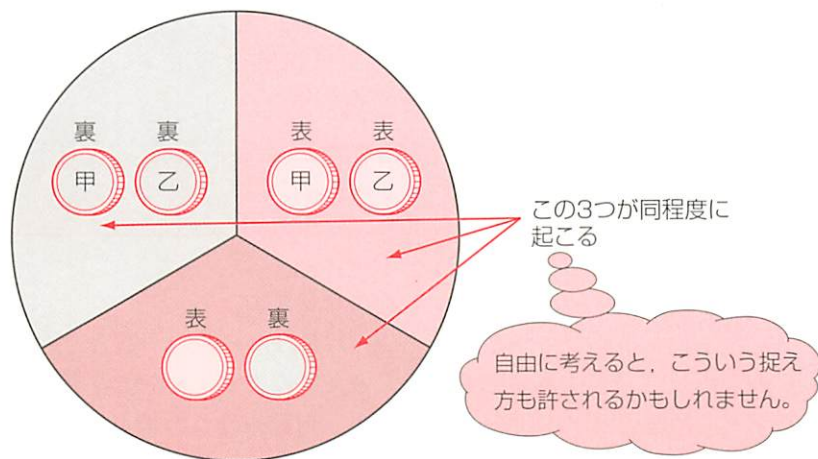
とは少し違うようです。したがって、ダランベールの考えを支持することはできません。彼は2つの事象{甲表、乙裏}、{甲裏、乙表}をひとまとめにして{一方が表、他方が裏}と捉えました。ここまでは誤りではないのですが、この事象の起こり方が、他の2つの事象{甲表、乙表}、{甲裏、乙裏}と同等だと見なしてしまいました。ここに誤りがあるようです。現象の表面的なところで判断すると危険であるという例です。正しくは、{甲表、乙表}、{甲表、乙裏}、{甲裏、乙表}、{甲裏、乙裏}の4通りが同程度に起こると考えて「少なくとも一方が表」である確率は、

$$\frac{3}{4} = 0.75$$

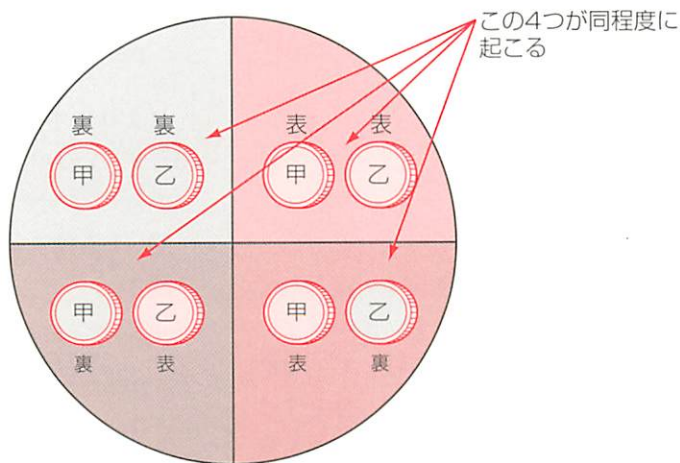
とすべきでしょう。標本空間をしっかりと見抜くことが大事なのです。

確率では標本空間の捉え方が大事

1 ダランベールの捉え方



2 通常の捉え方



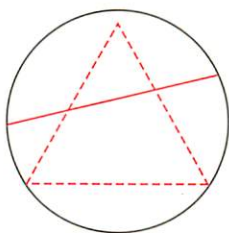
注) ダランベール (1717～1783)
フランスの啓蒙思想家、著名な数学者。百科全書派の重要な人物。

ここが要点!

ラプラス流の確率の考え方からすると、「同程度に確からしい」という前提さえ正しく捉えれば、同じ確率問題に対しては同じ答えが導き出されると考えられていました。ところが、1889年にベルトラン(Bertrand)という人が、「同程度に確からしい」ということの解釈によっては異なる確率が考えられる例を作りました。この例が以下に述べるベルトランの逆説というものです。

やさしい解説

「与えられた円に任意に1本の弦を引くとき、この弦の長さが内接正三角形の1辺の長さより大きくなる確率を求めよ」という問いに対して、ベルトランは次のように複数の異なる確率が考えられることを見つけました。



- (1) 対称性から弦の方向は一定として考えてもいいはず。その一定方向に垂直な直径ABに対して、中心OとAO、BOの中点をC、Dとするとき、弦がCDと交われば条件を満たすことがわかります(右ページ図1)。したがって、求める確率 p は、

$$p = \frac{\text{CDの長さ}}{\text{ABの長さ}} = \frac{1}{2}$$

となります。

- (2) 対称性から弦の一端Aを固定して考えてみます。Aを頂点とする内接正三角形の残りの頂点をB、Cとする。Aを一端とする弦AXの点Xが劣弧BC上にあれば、この弦AXは条件を満たします(右ページ図2)。したがって、求める確率 p は、

$$p = \frac{\text{劣弧BCの長さ}}{\text{全円周の長さ}} = \frac{1}{3}$$

となります。

- (3) 弦XYの中点Mが与えられた円の半径の半分の円にあれば、この弦XYは条件を満たします(右ページ図3)。

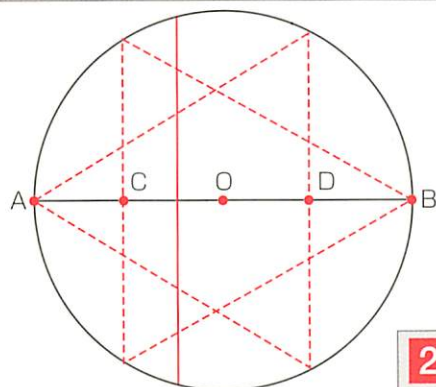
したがって、求める確率 p は、

$$p = \frac{\text{半径が半分の円の面積}}{\text{もとの円の面積}} = \frac{1}{4}$$

となります。

「同程度に確からしい」ということの解釈によって異なる確率

1 弦の方向に着目

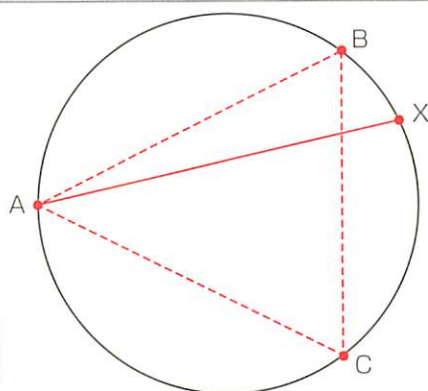


[図1]

この方向で考える

確率 $\frac{1}{2}$

2 弦の一端に着目

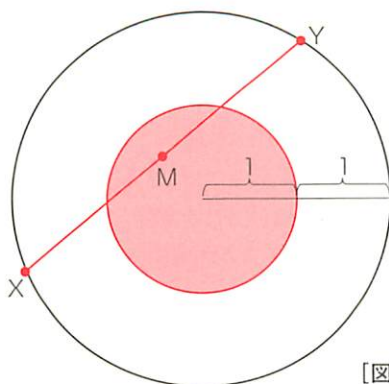


[図2]

確率 $\frac{1}{3}$

確率 $\frac{1}{4}$

3 弦の midpoint に着目



[図3]

ここが要点！

通常、われわれは体験したことを参考に、なんらかの予測を立てることが少なくありません。当然、体験する前とはその予測の確信度が異なってきます。確率の世界でも、起きてしまった事柄をもとに、その原因となる事柄の確率がどのくらいかを探る考えがあります。このとき使われるのがベイズの定理と呼ばれるものです。

やさしい解説

袋の中に4個の球あり、その中に白球が何個入っているかわかっていません。いま、1個の球を取り出したところ白球であったとします。このとき、袋の中に白球が3個入っている確率はどのくらいでしょうか。

原因は、次の4つの状態の袋であると考えられます。

A_1 …… 白1個, A_2 …… 白2個

A_3 …… 白3個, A_4 …… 白4個

これら4つの袋の存在確率は等しいものとしましょう。

A_1	A_2	A_3	A_4
$(A_1, ①)$	$(A_2, ①)$	$(A_3, ①)$	$(A_4, ①)$
$(A_1, ②)$	$(A_2, ②)$	$(A_3, ②)$	$(A_4, ②)$
$(A_1, ③)$	$(A_2, ③)$	$(A_3, ③)$	$(A_4, ③)$
$(A_1, ④)$	$(A_2, ④)$	$(A_3, ④)$	$(A_4, ④)$

次に、袋から1個の球を取り出したとき、それが白であるという事象を B とすると、次ページの図で見ると事象

B の要素の数が10個あり、そのうち3個が事象 A_3 の要素であることがわかります。したがって、袋の中に白球が3個入っている確率は事象 B が起きたときに事象が起こる確率 A_3 と考えられるので、求める答えは、

$$P_B(A_3) = \frac{P(A_3 \cap B)}{P(B)} = \frac{3}{10}$$

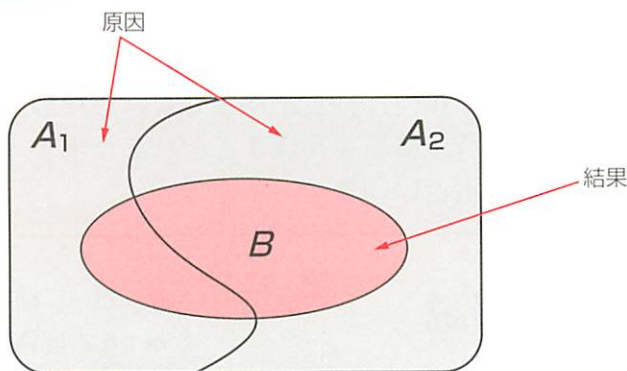
となります。

				B
$(A_1, ①)$	$(A_2, ①)$	$(A_3, ①)$	$(A_4, ①)$	
$(A_1, ②)$	$(A_2, ②)$	$(A_3, ②)$	$(A_4, ②)$	
$(A_1, ③)$	$(A_2, ③)$	$(A_3, ③)$	$(A_4, ③)$	
$(A_1, ④)$	$(A_2, ④)$	$(A_3, ④)$	$(A_4, ④)$	

一般に、事象 B が起きたときに事象 A が起こる確率(条件付き確率) $P_B(A)$ を「事象 B が起きたときにその原因が事象 A である確率」と解釈します。

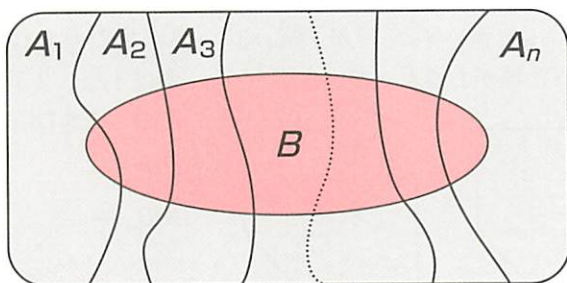
これが**ベイズの定理**で、 $P_B(A)$ のことを**原因の確率**と呼んでいます。

ベイズの定理の土台は条件付き確率



条件付き確率 $P_B(A_1) = \frac{P(A_1 \cap B)}{P(B)}$ を事象 B が起きたときに、その原因が A_1 である確率
 あると解釈するのがベイズの定理。 $P_B(A_2)$ も同様。

一般に原因がたくさんあるときは



$$P_B(A_i) = \frac{P(A_i) P_{A_i}(B)}{P(B)} = \frac{P(A_i) P_{A_i}(B)}{P(A_1) P_{A_1}(B) + P(A_2) P_{A_2}(B) + \cdots + P(A_n) P_{A_n}(B)}$$

を事象 B が起きたときに、その原因が A_i である確率と解釈する。

ただし

$$U = A_1 \cup A_2 \cup \cdots \cup A_i \cup \cdots \cup A_n,$$

$$A_i \cap A_j = \phi \quad (i \neq j), \quad A_i \neq \phi \quad (i = 1, 2, 3, \dots, n)$$

ここが要点！

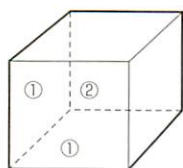
コインを投げたときの表裏の出方は、次にコインを投げた場合の表裏の出方に影響しません。試行が独立だからです。しかし、過去の試行の結果が、次の試行に影響を及ぼすこともあります。このようなものをマルコフ過程といいます。

やさしい解説

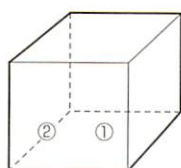
過去何回の試行の結果が、次の試行の結果に影響を及ぼすかによって、マルコフ過程は**単純（一重）マルコフ過程**、**二重マルコフ過程**、……、 **n 重マルコフ過程**と呼ばれます。

ここでは単純マルコフ過程の例をもとに、マルコフ過程を調べてみましょう。

(例) 箱Ⅰには1と書かれた球が2個、2と書かれた球が1個入っています。箱Ⅱには1と書かれた球が1個、2と書かれた球が1個入っています。



箱Ⅰ



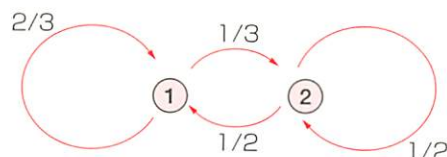
箱Ⅱ

す。以下、同様にして、記録した数字が1ならば箱Ⅰから、2ならば箱Ⅱから球をデタラメに1個選び、その数字を記録します。このようにして、1と2の数値の系列が得られます。

1, 2, 1, 2, 1, 2, 1, 1, 1, 2, 2, 2, 2, ……

そしてこの場合、途中において1が出るか2が出るかの確率は、その直前の結果のみで決定されます。したがって、この系列はマルコフ過程です。

1が出たときは、その次に1が出る確率は2/3、2が出る確率は1/3です。2が出たときは、その次に1が出る確率は1/2、2が出る確率は1/2です。このことを図示すると、次のようになります。



この図を**推移図（シャノン線図）**といい、確率を考える上で便利です。

初めにどちらか一方の箱からデタラメに球を1個選び、その球に書かれた数字を記録します。次に記録した数字が1ならば箱Ⅰから、2ならば箱Ⅱから球を1個選び、その数字を記録しま

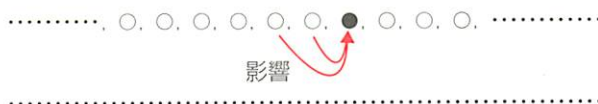
マルコフ過程の確率関係は推移図で

1 過去の影響回数によるいろいろなマルコフ過程

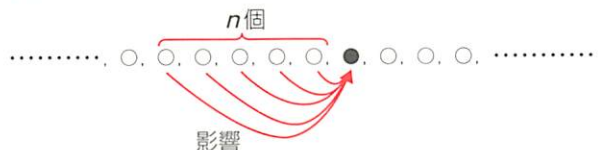
一重(単純)マルコフ過程



二重(単純)マルコフ過程

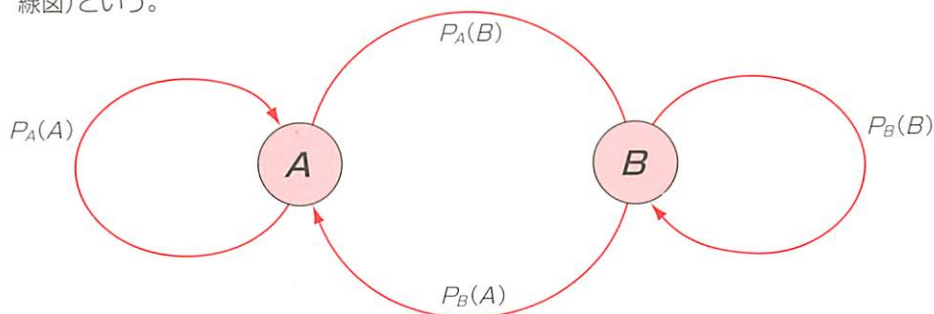


n 重(単純)マルコフ過程



2 推移図(シャノン線図)

マルコフ過程で、どんな確率で状態が推移するのかを図示したものを推移図(シャノン線図)という。



ここで、 $P_A(A)$ は、状態Aから状態Aに推移する確率
 $P_A(B)$ は、状態Aから状態Bに推移する確率
 $P_B(A)$ は、状態Bから状態Aに推移する確率
 $P_B(B)$ は、状態Bから状態Bに推移する確率

§ 26 期待値を知らずして賭をすべからず

ここが要点!

賭事をするときの費用や労力などをリスクといいます。これに対して見返りをリターンといいます。このリターンは、統計学では期待値(期待金額)の考えにつながります。この期待値を知らずして、いっぱしのギャンブラー気取りは笑止千万です。

やさしい解説

下表のような総本数100本の宝くじがあったとします。あなただったら、このくじ、いくらだったら買いますか。まさしく、リスクとリターンを考慮した意志決定です。

	1等	2等	はずれ
賞金	1000円	100円	0円
本数	1本	10本	89本
当たる確率	1/100	10/100	89/100

ここで、賞金総額をくじの総本数で割ったものを考えます。この場合、

$$\frac{1000 \times 1 + 100 \times 10 + 0 \times 89}{100} = 20$$

となります。この値をくじの**期待値**(または**期待金額**)といいます。

これは、平均すればくじ1本につき20円当たることが期待されるというわけです。または、くじ1本につき20円払いもどされるとも考えられます。

したがって、くじ1本引くのに20円未満であればドンドン買うべきです。ちょうど20円であればどちらで

もいいでしょう。20円を超えたら買ってはいけません。たとえば、このくじを1枚100円で買ったなら平均すると20円ぐらいしかもどってこないから80円の損ということになります。

また、もし1枚を10円で買えるなら全部のくじを買い占めるべきです。このように、いくら払ったら確率的にいくらもどってくるのかがはっきりしていないと、怖くて賭などできません。

それでは、実際のギャンブルの世界での期待金額を紹介しましょう。ただし、ここでは比較するために期待金額そのものではなく平均リターンを掲載します。

ギャンブルの種類	平均リターン
レース関係	約75%
パチンコ関係	約90%
くじ関係	約50%

ここで、レース関係には競馬、競輪、競艇、オートレースなどがあり、パチンコ関係にはパチンコ、パチスロ、くじ関係には宝くじ、toto、ロト6、ナンバーズなどがあります。

宝くじ・競輪・競馬・パチンコ, どの人がギャンブラー

1 宝くじの払戻金は 50 %以下!!

あなたはそれでも買うか???
「宝くじ買えば買うほど損をする。
だけど、買わなきゃ当たらない」

<1ユニット1000万本の当たりくじ>

1等	2億円	2本
1等の前後賞	5000万円	4本
1等の組違い賞	10万円	100本
2等	1000万円	2本
3等	100万円	40本
4等	1万円	1万本
.....		
.....		

宝くじを全部買
い占めても半分
以下しかもどら
ない。

2 競輪・競馬の払戻金は 75 %!!

大数の法則によると賭ける回数がふえるほど
トータルの赤字は25%に近づく。
「ギャンブルは回数を減らして運にまかせろ」

馬や騎手など誰の
お金で養うの



2 パチンコは 90 %!!



どんな日でもパチンコ屋は大にぎわい。
「10%の赤字ならマイイッパ」
でも、本当はもっと損をしているみたい!!

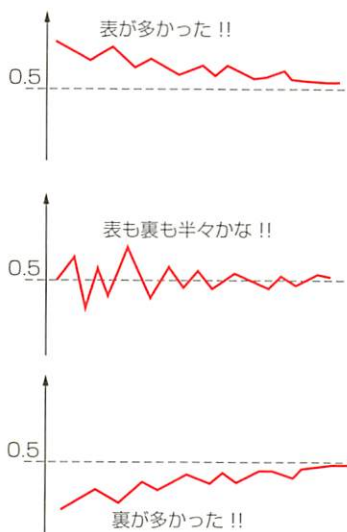


ここが要点！

確率は無限回の試行を前提としています。したがって、数少ない試行では偏りが生じます。コインを投げて表が出る確率が0.5でも10回投げて7回表であることは珍しくありません。もし、人生における賭がくじ引きのようなものであれば、ついている人・いない人がいても当たり前です。

やさしい解説

表が出る確率が0.5の硬貨を投げて、 n 回目までに表が出た相対度数を r とします。 n が小さいときは r の値はいろいろな変動をしますが、 n が大きくなっていくと r の値は例外なく0.5に近づいていきます。



この例からわかるように、人生が無限でコインの表裏のような賭を無数にできれば、誰もが運・不運は同じになります。しかし、人生有限なため、ついている人・いない人が存在します。

なお、確率現象のこのような性質は**大数の法則**と呼ばれるもので、一般には次のようにまとめられます。

「母平均 m の母集団から大きさ n の標本を無作為に抽出するとき、その標本平均 $\bar{X}$ は、 n が大きくなるに従っていくらでも母平均 m の近くに分布する」と。

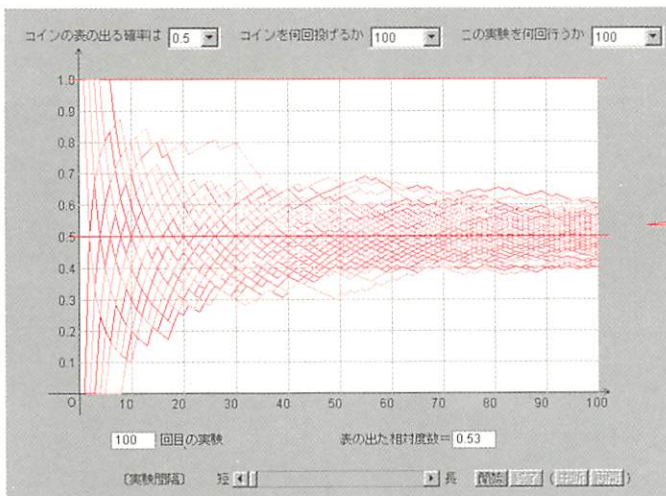
火事の件数、船の沈没による被害などは、昔は「運が悪かった」で終わっていましたが、統計データをたくさん集めた結果、「一定の法則」が見つかるようになりました。

「1年間に300件あたり1件の割合で火事による被害が起き、その被害額は平均100万円……」などです。このような法則が大数の法則で、火災保険、海上保険、自動車保険などの保険の基本となるものです。

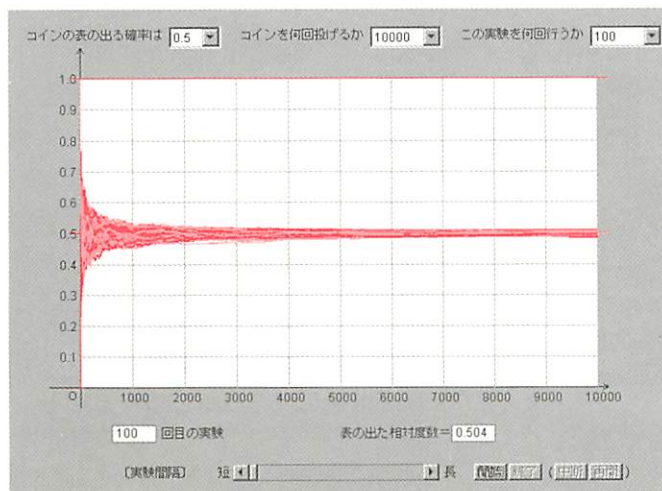
人生の裏に「大数の法則」あり

コインを n 回投げて、表の出た相対度数の変化をコンピュータでシミュレートしてみよう。

1 $n = 100$ として 100 回実験してみた



2 $n = 10000$ として 100 回実験してみた



§ 28 貧乏人が賭で金持ちに勝つ方法

ここが要点!

賭事の勝ち負けは勝率が同じであれば、所有する賭金の量に比例することがわかっています。つまり、貧乏人は金持ちに勝てないのです。しかし、勝率が0.5を少しでも上回れば事情は変わります。貧乏人でも、金持ちを打ち負かすチャンスは格段に向上します。

やさしい解説

ここでは、賭に勝つ能力と最初の財産(賭金)の量が賭事の最終的な勝敗にどう影響するのか、次の例で調べてみます。

A, B 2人は、次のルールで賭をすることにします。つまり、1回の賭で負けた方は勝った方に1万円を上げることにします。この賭を続けてA, Bのどちらか一方が破産したときに賭をやめることにします。



A, B両方とも勝つ確率が0.5であれば、Aが a 万円, Bが b 万円持っていれば、AがBを破産させる確率 P_A は次のようになります。

$$P_A = \frac{a}{a+b} \dots\dots\dots ①$$

BがAを破産させる確率 P_B は、

$$P_B = \frac{b}{a+b}$$

となります。つまり、賭の勝敗は、最初に所有している賭金に比例するのです。

ところが、A, Bの勝率が0.5でない場合には事情が変化します。Aの勝率を p , Bの勝率を $q (=1-p)$ とすると、

$$P_A = \frac{1 - \left(\frac{q}{p}\right)^a}{1 - \left(\frac{q}{p}\right)^{a+b}} \dots\dots\dots ②$$

$$P_B = \frac{1 - \left(\frac{p}{q}\right)^b}{1 - \left(\frac{p}{q}\right)^{a+b}}$$

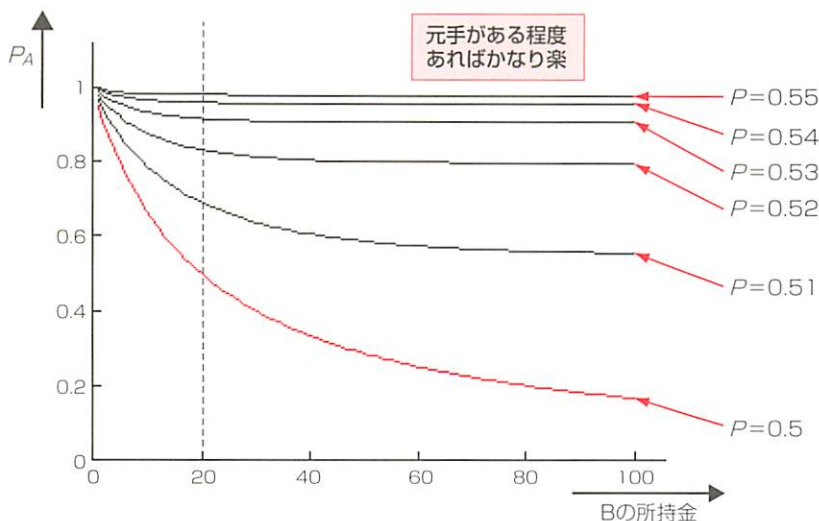
となります。式①, ②の確率がどういふものであるかは、次ページのグラフを見てください。

いろいろな不利な条件を抱えていても、能力を磨き、多数回チャレンジすれば世界が開けるかもしれません。

資金力の不足は能力で、でも元手は欲しい

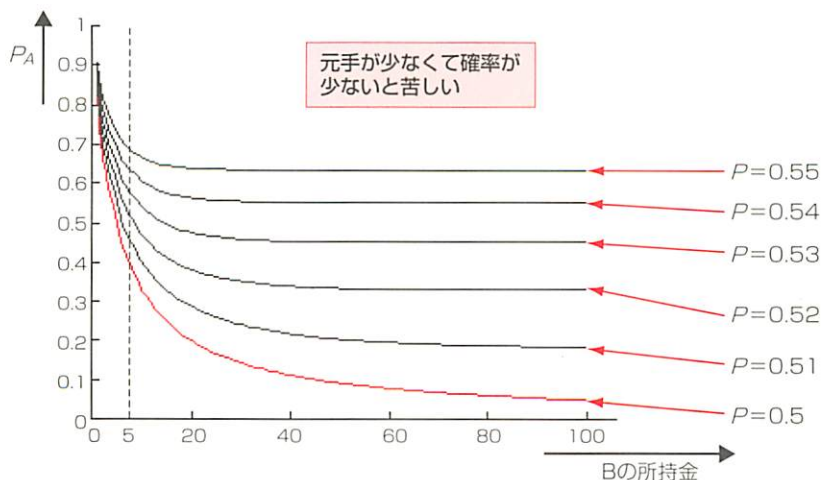
1 Aが20万円、Aの勝率 P_A が0.5～0.55のときにBを破産させる確率

ただし、Bの所持金は1～100万円、1回の賭金は1万円。



2 Aが5万円、Aの勝率 P_A が0.5～0.55のときにBを破産させる確率

ただし、Bの所持金は1～100万円、1回の賭金は1万円。



ここが要点!

コインやサイコロを使った賭博は、インチキでもない限り理論的には儲けることができません。しかし、トランプや花札、麻雀での賭博となると話は多少違ってきます。配られたカードや「パイ」は偶然でも、その後の扱いは人間の判断が入り込むからです。この結果、強い人や弱い人がいるわけです。そして、勝率が0.5を少しでも上回れば、たくさん賭をすることによって富を得ることができます。

やさしい解説

次のモデルで調べてみましょう。

A, Bの2人がゲームをし、Aが勝ったらBから1000円もらい、Aが負けたらBに1000円払うことにします。ただし、AのBに対する勝率を0.6とします。このゲームを100回行ったときのAの期待金額 E を調べてみましょう。

100回中 k 回Aが勝つ確率 P_k は、次のようになります(独立試行の定理)。

$$P_k = {}_{100}C_k (0.6)^k (1-0.6)^{100-k} \dots \textcircled{1}$$

ただし、 $k=0, 1, 2, \dots, 100$

また、このときのAさんの利益は、
 $1000k - 1000(100 - k)$

$$= 2000k - 1000 \times 100 \dots \textcircled{2}$$

となります。

したがって、Aの期待金額 E は $k=0, 1, 2, \dots, 100$ と変化したときの式①と②の積の総和ですから、次のようになります。

$$\begin{aligned} E &= \sum_{k=0}^{100} (2000k - 1000 \times 100) P_k \\ &= 2000 \sum_{k=0}^{100} k P_k - 1000 \\ &\quad \times 100 \sum_{k=0}^{100} P_k \end{aligned}$$

ここで、式①の確率分布は二項分布なので、

$$\sum_{k=0}^{100} k P_k = 100 \times 0.6, \quad \sum_{k=0}^{100} P_k = 1$$

したがって、

$$\begin{aligned} E &= 2000 \times 100 \times 0.6 - 1000 \times 100 \times 1 \\ &= 120000 - 100000 = 20000 \end{aligned}$$

となります。いかがでしょうか。

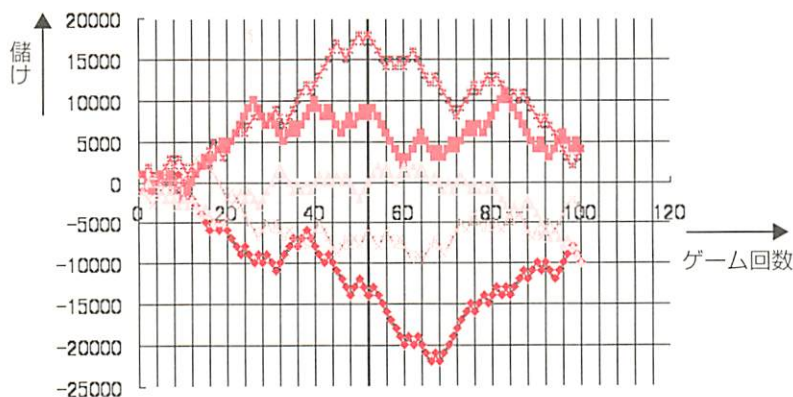
以上をゲームの回数を n 、Aの勝率を p 、1回の賭金を s としてまとめると、期待金額 E は次のようになります。

$$\begin{aligned} E &= \sum_{k=0}^n (2 \times s \times k - s \times n) P_k \\ &= 2 \times s \sum_{k=0}^n k P_k - s \times n \sum_{k=0}^n P_k \\ &= 2 \times s \times n \times p - s \times n \\ &= (2 \times s \times p - s) \times n \dots \dots n \text{の1次式} \end{aligned}$$

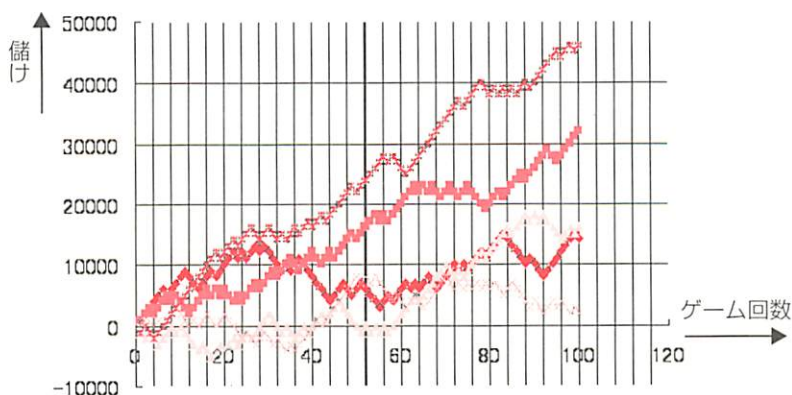
賭に自身があれば何回も何回も賭けろ

このことは次のシミュレーションで一目瞭然

- 1** 下図はゲームの回数を100、Aの勝率を0.5、1回の賭金を1000円として利益の推移を5回シミュレーションした例である。儲けの期待金額は0円とする。



- 2** 下図はゲームの回数を100、Aの勝率を0.6、1回の賭金を1000円として利益の推移を5回シミュレーションした例である。儲けの期待金額は20,000円とする。



§ 30 「あみだくじ」に落とし穴

ここが要点!

「あみだくじ」は、多くの人に親しまれています。くじを引く人数分だけ縦線を引き、これらをランダムに何本かの横線で結び、上下の対応関係をデタラメにチェンジしたものです。しかしこのとき、上下の対応関係は同等ではありません。上の部分は、その真下の部分に対応しやすくなっているのです。

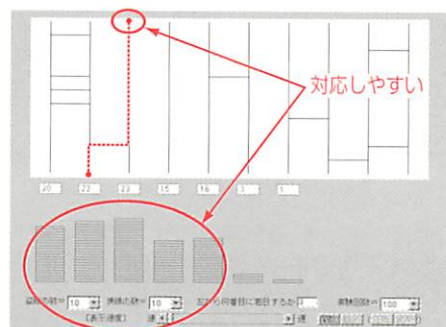
注) 無数に横線を引けば、上下の対応関係は同等になります。

やさしい解説

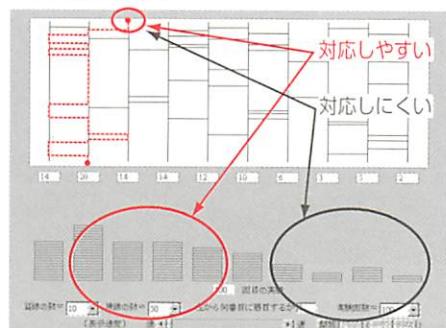
日常生活において何か決め事をするとき、ジャンケンやコイン、あみだくじなどがよく利用されます。それは多分、これらの手段が手軽で確率的にも平等だと思われるからでしょう。しかし、コインはともかく、ジャンケンについてはグウ、チョキ、パーのおおのの出方が確率的に同等かどうかは疑問です。それに、前後の出方は互いに影響合っているようにも思えます。そのため、ジャンケンを嫌う人も少なくありません。

それでは、あみだくじはどうでしょうか。このくじは、ジャンケンよりも確率的に平等だと思っている人が少なくありません。しかし、このくじは慎重に扱わないと不公平なくじになってしまいます。なぜならば、上下の対応関係は同等ではないからです。このことは、次のシミュレーション例から明らかです。

下図は、くじの数が10本で横線が10本のあみだくじの左から3番目がどこに対応しやすいかを、あみだくじを100個コンピュータで作成して、度数分布を調べた例です。

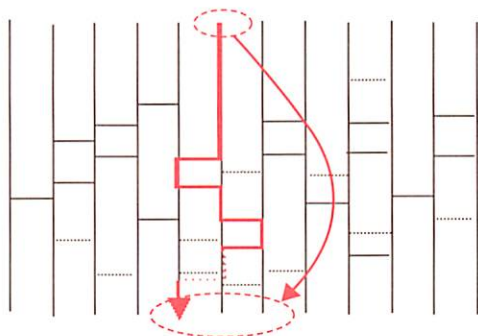


下図は、上の場合で横線が100本にした場合です。

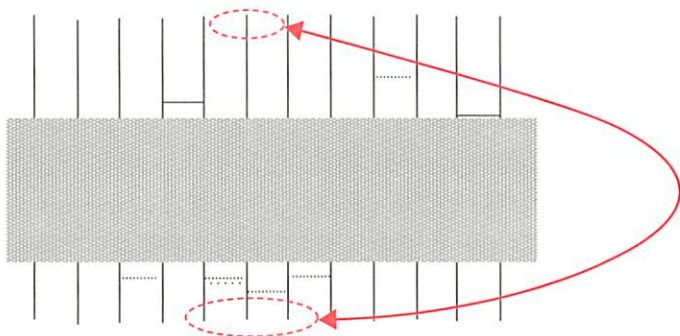


あみだくじ，上下の対応に偏りが

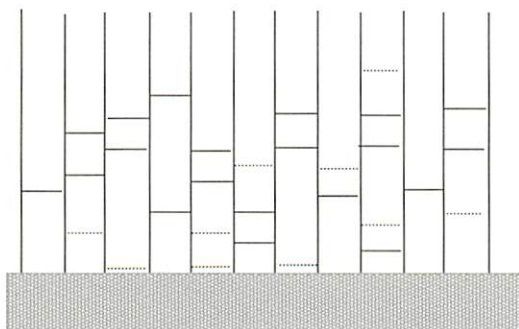
1 あみだくじは真下付近に対応しやすい。



2 途中を隠しても **1** の性質が活きる。



3 下部が隠されていれば公平なくじ。しかし，あみだくじでなくてもくじになる。



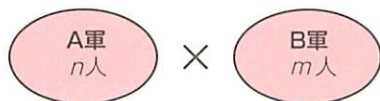
ここが要点！

ここが要点！ 昔からの言葉に「多勢に無勢」(人数に差がありすぎて勝ち目がない)とか、「衆寡適せず」というものがあります。しかし、少数でも精鋭であれば多くの相手を打ち負かすことができるかもしれません。つまり、「量より質」かもしれません。どちらが正しいか確率論で調べてみましょう。

やさしい解説

「多勢に無勢」を次のモデルを使って調べてみましょう。

いま、ここに2つの軍隊A, Bがあ
って互いに戦っているものとします。
A軍隊の兵士同士は個々同じ戦闘能力
を、またB軍隊の兵士同士も個々同じ
戦闘能力を持っているものとします。
そして、A軍隊の1人の兵士がB軍隊
の1人の兵士を打ち負かす確率(戦闘
能力)を p , B軍の兵士がA軍の兵士
を打ち負かす確率を q とします。ただ
し、 $p+q=1$ です。また、戦いは必
ず1対1で行われ、個々の兵士は何回
勝っても戦闘能力は衰えないものと
します。そして、A軍隊の兵士の数は n
人、B軍隊の兵士の数は m 人とします。



このとき、A軍の勝つ確率は次のように計算されます。

m 回の対戦でA軍が勝つ確率は

$${}_{m-1}C_{m-1}p^m$$

a a a a a a a a a a a a a

$m-1$ 回中、 $m-1$ 回A軍の兵士が勝つ
 m 回目もA軍の兵士が勝つ

$m+1$ 回の対戦でA軍が勝つ確率は、 ${}_m C_{m-1} p^m q^1$

a a a a a a a a b a a

m 回中、 $m-1$ 回A軍の兵士が勝つ
 $m+1$ 回目もA軍の兵士が勝つ

同様に、 $m+n-1$ 回の対戦でA軍が勝つ確率は、 ${}_{m+n-2}C_{m-1}p^mq^{n-1}$

よって、A軍が勝つ確率は以上の確率をすべて加えて、

$$\begin{aligned} & {}_{m-1}C_{m-1}p^m + {}_mC_{m-1}p^mq^1 \\ & + {}_{m+1}C_{m-1}p^mq^2 + {}_{m+2}C_{m-1}p^mq^3 \\ & + \cdots + {}_{m+n-2}C_{m-1}p^mq^{n-1} \cdots \textcircled{1} \end{aligned}$$

また、同様にしてB軍が勝つ確率は、

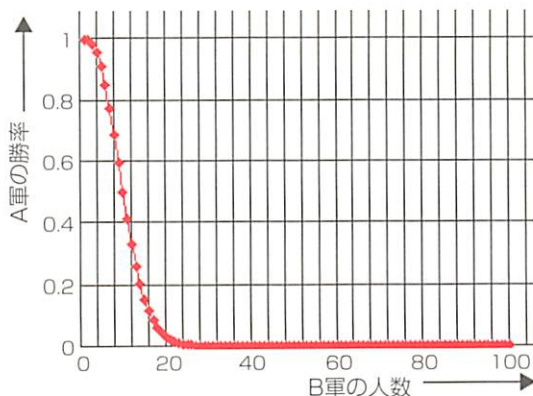
$$\begin{aligned} & {}_{n-1}C_{n-1}q^n + {}_nC_{n-1}pq^n \\ & + {}_{n+1}C_{n-1}p^2q^n + {}_{n+2}C_{n-1}p^3q^n \\ & + \cdots + {}_{n+m-2}C_{n-1}p^{m-1}q^n \quad \cdots (2) \end{aligned}$$

式①、②ではピンときません。具体的な数値で調べたのが、次ページのグラフです。

「多勢に無勢」, 「量より質」どっちもどっち

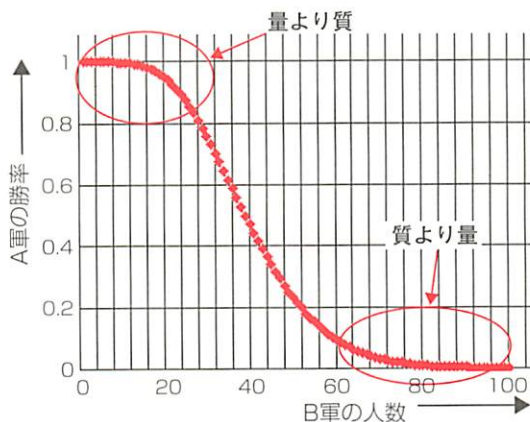
1 A 軍が勝つ確率

A軍：能力 0.5
人数 10
B軍：能力 0.5
人数 1~100



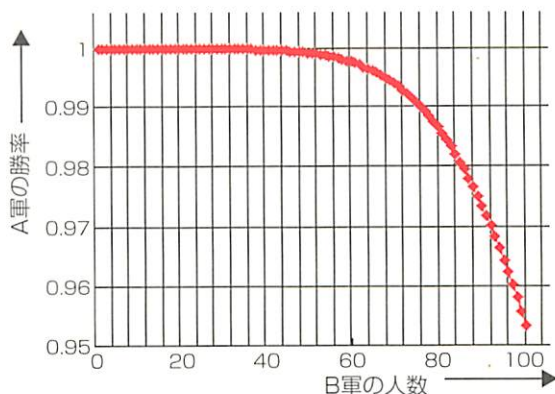
2 A 軍が勝つ確率

A軍：能力 0.8
人数 10
B軍：能力 0.2
人数 1~100



3 A 軍が勝つ確率

A軍：能力 0.95
人数 10
B軍：能力 0.05
人数 1~100



ここが要点!

「次のことは正しいか、○か×をつけて答えなさい」という試験が○×試験です。別名「2 択問題」ともいわれています。この試験は、勉強嫌いな学生にはすごくありがたい試験なのです。なぜならば、鉛筆を転がすだけで 50 点ぐらいはとれてしまうからです。さすが「3 択問題」になるところはいきません。しかし、この場合でも平均すると 30 点以上とれてしまいます。一般に「 n 択問題」だと、100 点満点の 100 を n で割った数が期待得点になります。

やさしい解説

昔から勉強の嫌いな学生さんはたくさんいました。でも、試験だけは合格しないと、親が出してくれた高い学費は無駄になってしまいます。こんな彼らは、○×試験に出会うと涙が出ます。なぜなら、まったく勉強していなくても普通の「つき」で満点の半分くらいの点数が確保できるからです。このことを、「○×問題が 10 問あって、各 10 点で 100 点満点」の試験で調べてみましょう。

この試験にデタラメに答えると 1 問正解である確率は $1/2$ 。したがって、10 問中 k 問解ける確率 P_k は独立試行の定理から、

$$\begin{aligned} P_k &= {}_{10}C_k \left(\frac{1}{2}\right)^k \left(\frac{1}{2}\right)^{10-k} \\ &= {}_{10}C_k \left(\frac{1}{2}\right)^{10} \end{aligned}$$

具体的に数値を求めると、

$$k=0 \text{ のとき, } P_0 = 0.000977$$

$$k=1 \text{ のとき, } P_1 = 0.009766$$

$$k=2 \text{ のとき, } P_2 = 0.043945$$

$$k=3 \text{ のとき, } P_3 = 0.117188$$

$$k=4 \text{ のとき, } P_4 = 0.205078$$

$$k=5 \text{ のとき, } P_5 = 0.246094$$

$$k=6 \text{ のとき, } P_6 = 0.205078$$

$$k=7 \text{ のとき, } P_7 = 0.11718$$

$$k=8 \text{ のとき, } P_8 = 0.043945$$

$$k=9 \text{ のとき, } P_9 = 0.009766$$

$$k=10 \text{ のとき, } P_{10} = 0.000977$$

となります。したがって、この試験を受けてデタラメに答えたときの正解数の期待値 E は、次のようになります。

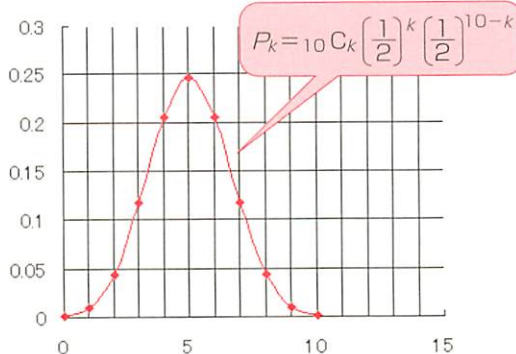
$$\begin{aligned} E &= 0 \times P_0 + 1 \times P_1 + 2 \times P_2 + \cdots \\ &\quad \cdots 9 \times P_9 + 10 \times P_{10} \\ &= 5 \end{aligned}$$

どうでしょうか。デタラメに答えても、10 問中 5 問解けることが期待できるというわけです。

n 択問題の期待正解数は問題数 $\div n$

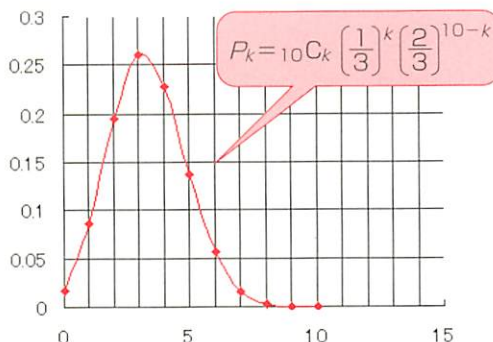
1 2 択問題

10問にデタラメに答えたときの正解数の確率分布。
期待値は5問



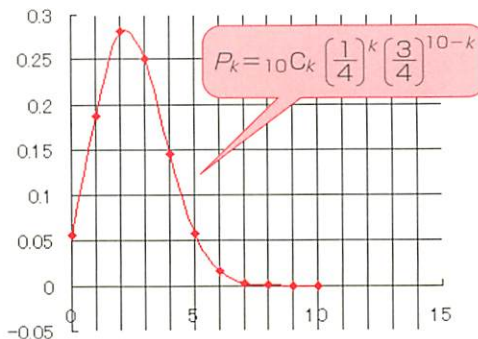
2 3 択問題

10問にデタラメに答えたときの正解数の確率分布。
期待値は3.333...問



3 4 択問題

10問にデタラメに答えたときの正解数の確率分布。
期待値は2.5問



ここが要点！

確率現象を捉えた諺(ことわざ)に表題のものがあります。これは経験から出てきたのでしょうか、数学として正しいことを説明できます。この諺の真意を正しく使うと、人生得することも少なくありません。

やさしい解説

ここに、1回の射撃で的(まと)に当たる確率が1/10の射手がいるとします。この射手が7回射撃したときに、少なくとも的に1回当たる確率はどのくらいになるのかを求めてみましょう。

実際の射撃では、ある回に当たるか当たらないかは、次の回の射撃に影響するようですが、ここでは各回の射撃は独立であるという数学のモデルで考えることにします。すると、排反事象と試行の独立の考え方から、この射手が7回射撃したとき、少なくとも1回当たる確率 P は次のようになります。

$$\begin{aligned} P &= 1 - (\text{7回とも当たらない確率}) \\ &= 1 - \left(1 - \frac{1}{10}\right)^7 \end{aligned}$$

この値を計算(電卓でもできそうです)すると、

$$P = 0.5217$$

となります。

あまり上手でなくても、何回も挑戦すると成功する確率がグンと増えているのがわかります。

この射手が7回ではなく100回射撃したとすると、少なくとも1回当たる確率 P は同様に考えて、

$$P = 1 - \left(1 - \frac{1}{10}\right)^{100}$$

となります。電卓ではちょっと大変なので、表計算ソフトのExcelで数値を求めてみたのが次の値です。

$$P = 0.99997$$

ほぼ1といってもいいでしょう。すごいですね。

ここで、この話を一般化してまとめておきましょう。

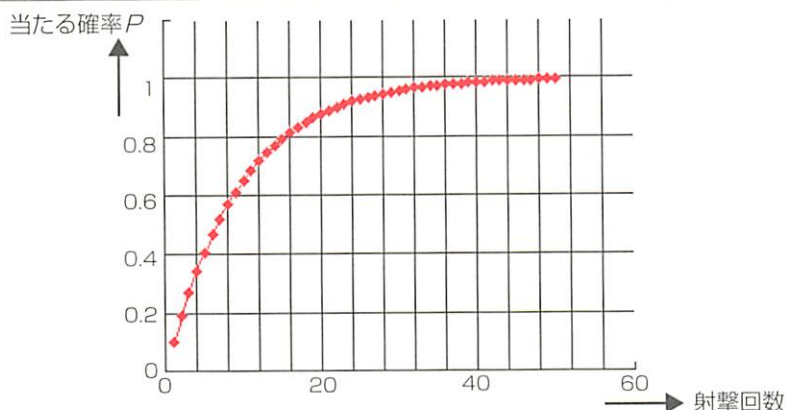
1回の射撃で的に当たる確率が r の射手がいるとします。この射手が n 回射撃したときに、少なくとも的に1回当たる確率 P は、

$$P = 1 - (1 - r)^n$$

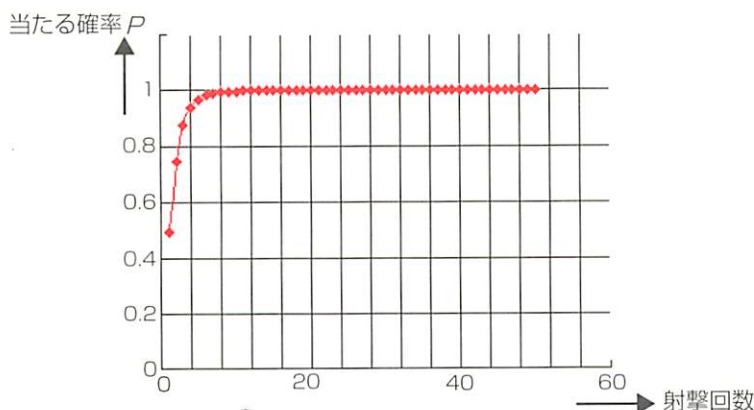
となります。

失敗しても諦めずに何回も挑戦しよう

- 1 1回の射撃で的に当たる確率が0.1の射手が n 回射撃したときに、少なくとも1回当たる確率 P



- 2 1回の射撃で的に当たる確率が0.5の射手が n 回射撃したときに、少なくとも1回当たる確率 P



絶望的だ!!

§ 34 くじ引きは先を争うことなかれ

ここが要点!

「くじ引き」の際に先を争って引く光景をよく目にします。しかし、くじ引きは引く順番によって当たる確率が変化するものなののでしょうか。幸か不幸か、確率の考えでは引く順番にはまったく関係ないようです。

やさしい解説

「くじ引き」において、当たる確率は引く順序に関係ないことを次の例で調べてみましょう。

5本のくじがあり、そのうち2本が当たりくじであるとしします。このとき、このくじから非復元抽出で1本のくじを3人が引くとき、1, 2, 3番目に引く人が当たる確率をおのおの求めてみましょう。

(1) 1番目に引く人の当たる確率

この確率は $\frac{2}{5}$ です。

(2) 2番目に引く人の当たる確率

2番目に引く人が当たるのは、次の2つの場合です。つまり、1番目の人が当たって2番目の人が当たる場合と、1番目の人がはずれて2番目の人が当たる場合です。したがって、2番目に引く人が当たる確率は、

$$\begin{aligned}\frac{2}{5} \times \frac{1}{4} + \frac{3}{5} \times \frac{2}{4} &= \frac{2+6}{20} \\ &= \frac{8}{20} = \frac{2}{5}\end{aligned}$$

(3) 3番目に引く人の当たる確率

3番目に引く人が当たるのは次の3つの場合です。つまり、1番目の人が当たり2番目の人がはずれ3番目が当たる場合、1番目の人がはずれ2番目の人が当たり3番目が当たる場合、1番目の人がはずれ2番目の人がはずれ3番目が当たる場合です。したがって、3番目に引く人が当たる確率は、

$$\begin{aligned}\frac{2}{5} \times \frac{3}{4} \times \frac{1}{3} + \frac{3}{5} \times \frac{2}{4} \times \frac{1}{3} \\ + \frac{3}{5} \times \frac{2}{4} \times \frac{2}{3} &= \frac{2}{5}\end{aligned}$$

したがって、(1)、(2)、(3)いずれの場合も当たる確率は同じです。

なお、前にくじを引いた人が、その結果をまだ引いてない人に知らせてしまえば、その後の人のくじに当たる確率は当然変化します。たとえば、最初のほうの人が全部当たりくじを引いてしまえば、最後のほうの人は引く前から当たらないことがわかってしまいます。ここでは、あくまでも引いたくじを一斉に開くものとします。