2001年度版 基礎数学ワークブック 番外編 No.1 「確率」

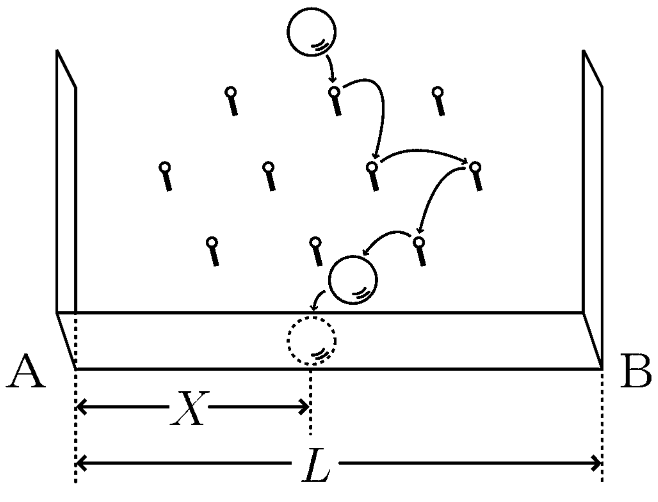
高知工科大学

基礎数学ワークブック

(2001年度版)

番外編 1

「確率」



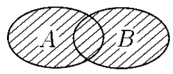
井上 昌昭 著

＜ 記号 ＞

このページでは本書で用いる数学記号をまとめる。

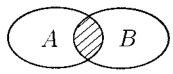
1.＜ 集合 ＞ 集合はアルファベットの大文字で表す。

$A \cup B$



（和集合）

$A \cap B$



（共通部分集合）

$A \bigcup B$ ： A と B の和集合

$A \bigcap B$ ： A と B の共通部分集合

ϕ ： 空集合

$[\ a\ ,\ b\ ] = \{\ x : a \leqq x \leqq b\ \}$ ：区間 (a 以上 b 以下の実数 x の全体)

$n(A)$ ：集合 A に含まれる要素の個数 (他の本では $\sharp A$ とか $|A|$ などと書く場合もある)

2.＜ 場合の数 ＞

${}_nP_r = n \times (n-1) \times \cdots \times (n-r+1)$ ：n 個のものから r 個とって一列に並べた順列の総数

$n! = {}_nP_n = n \times (n-1) \times \cdots \times 2 \times 1$ ：n の階乗 (ただし $0! = 1$ と定める)

${}_nC_r = \frac{{}_nP_r}{r!}$ ：n 個のものから r 個とった組合せの総数 (ただし ${}_nC_0 = 1$)

3.＜ 資料の整理 ＞ 実数値データは $x_1, x_2, \cdots, x_n$ などのようにアルファベットの小文字で表す。

m : 平均値 ，　v : 分散 ，　$\sigma = \sqrt{v}$: 標準偏差

4.＜ 確率 ＞

U ：全事象 (他の本では全事象を S , Ω 等で表す場合もある)

$P(A)$ ：事象 A の起こる確率

p , p_1 , p_2 , $\cdots$, p_n ：確率の値

X, Y ：確率変数

$P(X = a)$ ：$X = a$ となる確率

$P(a \leqq X \leqq b)$ ：X が a 以上 b 以下になる確率

$E[X]$ ：X の期待値 (平均)

$V(X)$ ：X の分散

$B(n, p)$ ：確率 p 次数 n の二項分布 $\left(P(X = k) = {}_nC_k\ p^k (1-p)^{n-k}\right)$

$N(m, v) = N(m, \sigma^2)$ ：平均 m 分散 v の正規分布 $\left(P(a \leqq X \leqq b) = \int_a^b \frac{1}{\sqrt{2\pi v}} e^{-\frac{(x-m)^2}{2v}} dx\right)$

5.＜ 本書だけの記号 ＞ 本書では「変数の標準化」をその変数に $*$ をつけて表現する。

① x^* ：数値データを標準化した値 (P7 参照)

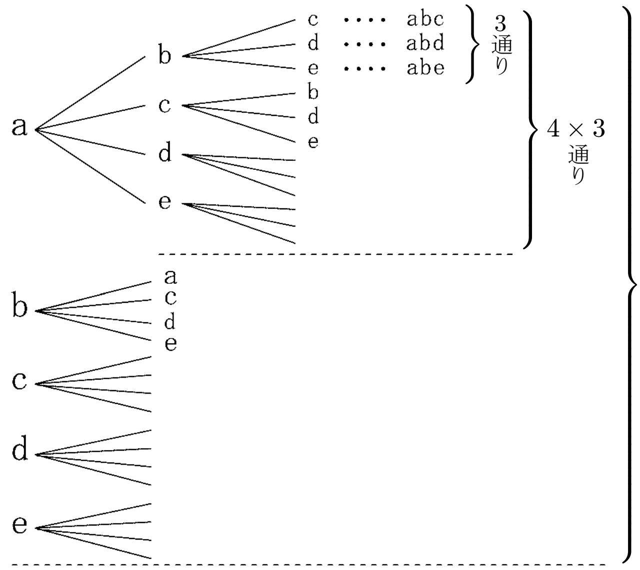
② $X^* = \dfrac{X - m}{\sigma}$ ：確率変数 X の標準化 (P23 , P34 参照)

この記号は便宜上つけたもので本書だけの記号である。気にいらない人は適当に記号をつけかえて読んでほしい。

＜順列＞

例 1 5 個のアルファベット a,b,c,d,e から 3 個えらんで 1 つの単語を作る。
3 文字で表される単語は何通りできるか? (ただし aab のように同じ文字を 2 回以上は使わない。)

(解) 1 文字目 2 文字目 3 文字目 単語

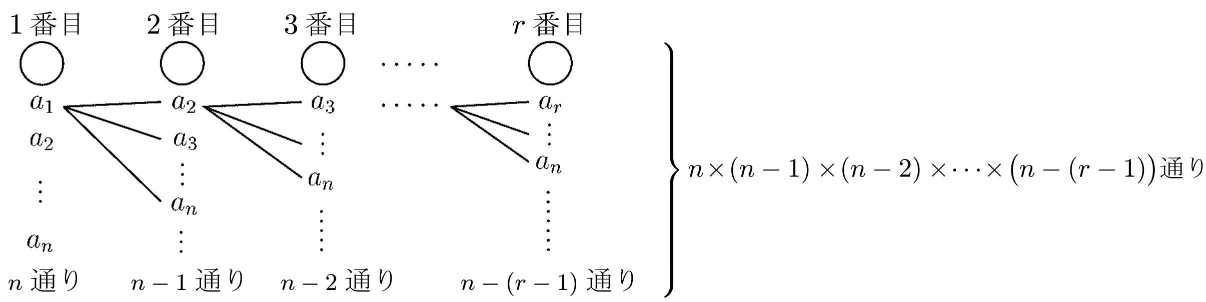


左の図 (樹形図という) から 1 文字目が 5 通り，2 文字目が 4 通り，3 文字目が 3 通りだから

(答) $5 \times 4 \times 3 = 60$ (通り)

(注) abc, abd などの単語を 5 個のものから 3 個とり出して一列に並べた**順列**という。この場合順列の総数は $5 \times 4 \times 3$ である。

一般に n 個の文字 $a_1, a_2, \cdots, a_n$ から r 個とり出して一列に並べた順列の総数を ${}_nP_r$ とすると，総数 ${}_nP_r$ は r 個の積で表される。



図より

$$\boxed{{}_nP_r = n \times (n-1) \times (n-2) \times \cdots \times (n-r+1)}$$ (n 個から r 個とった**順列**の総数)

また ${}_nP_n$ を $\boldsymbol{n}$ **の階乗**といい

$$\boxed{n! = {}_nP_n = n \times (n-1) \times (n-2) \times \cdots \times 3 \times 2 \times 1}$$ (**階乗**)

という記号で表す。

例 2 ${}_7P_2 = 7 \times 6 = 42$, ${}_6P_3 = 6 \times 5 \times 4$, $3! = 3 \times 2 \times 1 = 6$, $4! = 4 \times 3 \times 2 \times 1 = 24$

問 1 (1) ${}_{10}P_3 =$, (2) ${}_6P_4 =$, (3) $5! =$, (4) $6! =$

問 2 $1, 2, 3, 4$ の 4 個の数学を使って 3 桁の数を作る。以下の場合に 3 桁の数は何通りできるか?
(1) 同じ数字は 1 回しか使えない場合 (123〜432) (2) 同じ数字を何回使ってもよい場合 (111〜444)

＜組合せ＞

例 1 5 個のアルファベット a,b,c,d,e から 3 個えらんで 1 つの組を作る。このとき $\{a,b,c\}$，$\{a,b,d\}$，$\{a,b,e\}$，$\{a,c,d\}$，$\{a,c,e\}$，$\{a,d,e\}$，$\{b,c,d\}$，$\{b,c,e\}$，$\{b,d,e\}$，$\{c,d,e\}$

の計 10 組できる。一方並べる順も考えると以下のように 60 通りできる。

$$\left.\begin{array}{cccccccccc}
abc, & abd, & abe, & acd, & ace, & ade, & bcd, & bce, & bde, & cde \\
acb, & adb, & aeb, & adc, & aec, & aed, & bdc, & bec, & bed, & ced \\
bac, & bad, & bae, & cad, & cae, & dae, & cbd, & cbe, & dbe, & dce \\
bca, & bda, & bea, & cda, & cea, & dea, & cdb, & ceb, & deb, & dec \\
cab, & dab, & eab, & dac, & eac, & ead, & dbc, & ebc, & ebd, & ecd \\
cba, & dba, & eba, & dca, & eca, & eda, & dcb, & ecb, & edb, & edc
\end{array}\right\} 3! = 6\text{ 通り}$$

$${}_5P_3 \;=\; 5 \times 4 \times 3 \;=\; 60(\text{ 通り })$$

従って並べる順を考えない組の総数は

$$\frac{{}_5P_3}{3!} \;=\; \frac{5 \times 4 \times 3}{3 \times 2 \times 1} \;=\; 10(\text{組})$$

になる。

一般に n 個のものから r 個とり出して 1 つの組にしたものの総数を ${}_nC_r$ と書くと

$$\boxed{{}_nC_r \;=\; \frac{{}_nP_r}{r!} \;=\; \frac{n \times (n-1) \times \cdots \times (n-r+1)}{r \times (r-1) \times \cdots \times 1}} \quad \left(\begin{array}{l} n\text{ 個のものから } r\text{ 個とった} \\ \text{組合せの総数} \end{array}\right)$$

となる。

問 1 10 人から 4 人のリレー走者を選ぶ。

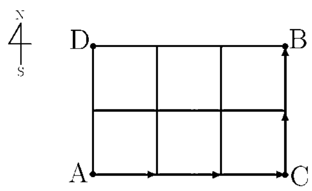
(1) 走る順番を考えると何通りできるか？

(2) 走る順を考えないで，ただ 4 人の組を作るときは何組できるか？

例 2 図 1 のような道がある町で A 地点から B 地点へいたる最短経路は何通りあるかを考える。
A から B へ行くには東へ 3 区画，北へ 2 区画進まねばならない。その経路は東 (3 個)，北 (2 個) の並びで表される。

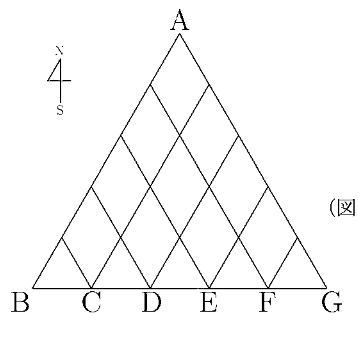
$${}_5C_3\text{通り}\left\{\begin{array}{llllll}
\text{東} & \text{東} & \text{東} & \text{北} & \text{北} & \cdots \quad \text{A} \to \text{C} \to \text{B の道順} \\
\text{東} & \text{東} & \text{北} & \text{東} & \text{北} & \cdots \\
 & & \vdots & & & \\
\text{北} & \text{北} & \text{東} & \text{東} & \text{東} & \cdots \quad \text{A} \to \text{D} \to \text{B の道順}
\end{array}\right.$$

よって最短経路は ${}_5C_3 \;=\; 10$ 通り。



(図1)

問 2 図 2 のような道で A 地点から出発し，南へ進む。B から G の各地点へいたる最短経路は何通りあるか。それぞれの地点について計算し，B から G の記号の下に " 何通り " かを記せ。



(図2)

＜ 二項定理 ＞

$(a+b)^n$ の展開式を計算する。

(パスカルの三角形)

$n=0$	$(a+b)^0=1$	1
$n=1$	$(a+b)^1=a\ +\ b$	1 1
$n=2$	$(a+b)^2=a^2\ +\ 2ab\ +\ b^2$	1 2 1
$n=3$	$(a+b)^3=a^3+3a^2b+3ab^2+b^3$	1 3 3 1
$n=4$	$(a+b)^4=a^4+4a^3b+6a^2b^2+4ab^3+b^4$	1 4 6 4 1
$n=5$	$(a+b)^5=a^5+5a^4b+10a^3b^2+10a^2b^3+5ab^4+b^5$	1 5 10 10 5 1
		$\| \ \| \ \| \ \| \ \| \ \|$
		${}_5C_0$ ${}_5C_1$ ${}_5C_2$ ${}_5C_3$ ${}_5C_4$ ${}_5C_5$

右図のように展開した各項の係数を三角形状に
並べたものをパスカルの三角形という。前ページ問 2 より，各係数は最短経路の場合の数と同じであるから

$$(a+b)^5 = {}_5C_0a^5b^0 + {}_5C_1a^4b^1 + {}_5C_2a^3b^2 + {}_5C_3a^2b^3 + {}_5C_4a^1b^4 + {}_5C_5a^0b^5$$

となる。一般に

$$\boxed{\begin{aligned}(a+b)^n &= {}_nC_0a^nb^0 + {}_nC_1a^{n-1}b^1 + {}_nC_2a^{n-2}b^2 + \cdots + {}_nC_{n-1}a^1b^{n-1} + {}_nC_na^0b^n \\ &= \sum_{k=0}^{n} {}_nC_ka^{n-k}b^k\end{aligned}}$$

となる。これを**二項定理**という。

(注) 組合せの意味から常に $\boxed{{}_nC_0 = {}_nC_n = 1}$ である。組合せを階乗で表現するために $\boxed{0! = 1}$ と定める。

問 1 ${}_nC_r = \dfrac{{}_nP_r}{r!}$ ， ${}_nP_r = \dfrac{n!}{(n-r)!}$ より ${}_nC_r = \dfrac{n!}{r!(n-r)!}$ である。

${}_nC_{n-r}$ を階乗の記号!を使って表せ。

問 2 $(a+b)^7$ の展開式を求めよ。

$(a+b)^7 =$

問 3 二項定理で $a=b=1$ とおくことにより，次の和を求めよ。

${}_nC_0 + {}_nC_1 + {}_nC_2 + \cdots + {}_nC_{n-1} + {}_nC_n =$

＜資料の整理 1＞

n 個の実数値データ $\{x_1, x_2, \cdots, x_n\}$ に対し

$$\boxed{m = \frac{1}{n}\{x_1 + x_2 + \cdots + x_n\} = \frac{1}{n}\sum_{k=1}^{n} x_k} \quad \text{(平均)}$$

を**平均**という。また平均 m に対し

$$\boxed{v = \frac{1}{n}\{(x_1 - m)^2 + (x_2 - m)^2 + \cdots + (x_n - m)^2\} = \frac{1}{n}\sum_{k=1}^{n}(x_k - m)^2} \quad \text{(分散)}$$

を**分散**といい，その正の平方根 $\sqrt{v}$ を**標準偏差**といい

$$\boxed{\sigma = \sqrt{v} = \sqrt{\frac{1}{n}\sum_{k=1}^{n}(x_k - m)^2}} \quad \text{(標準偏差)}$$

で表す。

例 1　7 個数値データ $\{0,\ 2,\ 4,\ 6,\ 8,\ 10,\ 12\}$ の平均 m は

$$m = \frac{1}{7}\{0 + 2 + 4 + 6 + 8 + 10 + 12\} = 6$$

分散は

$$v = \frac{1}{7}\{(0-6)^2 + (2-6)^2 + (4-6)^2 + (6-6)^2 + (8-6)^2 + (10-6)^2 + (12-6)^2\} = 16$$

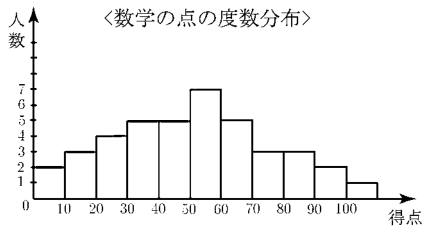
より標準偏差は $\sigma = \sqrt{v} = \sqrt{16} = 4$

問 1　以下の数値データの平均と分散および標準偏差を求めよ。

(1) $\{3,\ 4,\ 5,\ 6,\ 7,\ 8,\ 9\}$　　　(2) $\left\{-\frac{3}{2},\ -1,\ -\frac{1}{2},\ 0,\ \frac{1}{2},\ 1,\ \frac{3}{2}\right\}$

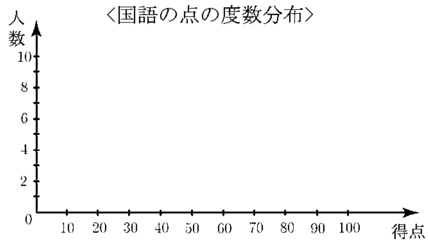
分散 (および標準偏差) は「データの広がりぐあい」を示す数値である。分散が小さいほど平均の近くにデータが集まっている。実際のデータの分布を表すには以下のようなグラフを使う。

例 2　以下のデータは 40 人の数学の試験結果を低い点から順に並べたものである。
{0, 6, 12, 15, 18, 20, 22, 25, 28, 30, 33, 35, 35, 38,
40, 40, 42, 45, 48, 50, 50, 52, 54, 55, 57, 58, 60,
60, 63, 65, 67, 70, 75, 77, 80, 84, 86, 90, 94, 100}



このテストの平均は 49.475，分散は 616.75，標準偏差は 24.83 である。上図に 0 点〜9 点が 2 人，10 点台が 3 人，$\cdots$，90 点台 2 人，100 点 1 人としたグラフを書いた。このようなグラフを**度数分布**または**ヒストグラム**という。

問 2　以下のデータは 40 人の国語の試験結果である。右図にヒストグラムを作り，平均・分散・標準偏差を求めよ。
{25, 28, 30, 32, 36, 38, 40, 42, 44, 46, 46, 48, 48, 50, 50 50,
52, 52, 54, 54, 56, 58, 60, 60, 62, 64, 66, 68, 68, 70, 72 74,
76, 78, 80, 80, 82, 84, 86, 90}
この問題にかぎり EXCEL や Mathematica 等を使ってよい。

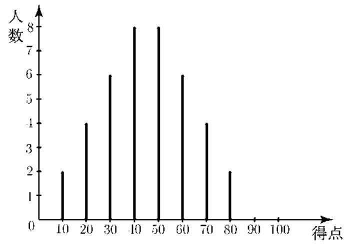


＜ 資料の整理 2 ＞

例 40 人のクラスで数学の試験を 2 回した。全て部分点なしであり，10 点きざみで点数がつけてある。このような場合に同じ点数が何人もいるので，以下の表のように結果を表す。またヒストグラムのかわりに右図のように棒グラフで表す。

(1 回目)

得点	0	10	20	30	40	50	60	70	80	90	100
人数	0	2	4	6	8	8	6	4	2	0	0



1 回目のテストの平均 m は

$$m = \frac{1}{40}\{10\times 2+20\times 4+30\times 6+40\times 8+50\times 8+60\times 6+70\times 4+80\times 2\} = 45$$

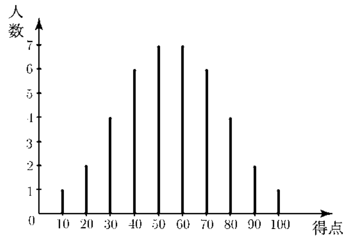
であり，分散 v は

$$v = \frac{1}{40}\{(10-45)^2\times 2+(20-45)^2\times 4+(30-45)^2\times 6+(40-45)^2\times 8$$
$$+(50-45)^2\times 8+(60-45)^2\times 6+(70-45)^2\times 4+(80-45)^2\times 2\} = 325$$

であり，標準偏差は $\sigma = \sqrt{v} = \sqrt{325} \fallingdotseq 18.03$ となる。

(2 回目)

得点	0	10	20	30	40	50	60	70	80	90	100
人数	0	1	2	4	6	7	7	6	4	2	1



2 回目のテストの平均 m，分散 v，標準偏差 σ は

$$m = \frac{1}{40}\{10\times 1+20\times 2+30\times 4+40\times 6+50\times 7+60\times 7+70\times 6+80\times 4+90\times 2+100\times 1\} = 55$$

$$v = \frac{1}{40}\{(10-55)^2\times 1+(20-55)^2\times 2+(30-55)^2\times 4+(40-55)^2\times 6+(50-55)^2\times 7$$
$$+(60-55)^2\times 7+(70-55)^2\times 6+(80-55)^2\times 4+(90-55)^2\times 2+(100-55)^2\times 1\} = 425$$

$\sigma = \sqrt{v} = \sqrt{425} \fallingdotseq 20.62$ となる。

この例のように同じ値がいくつかあるとき以下の形の表にする。

資料の値	x_1	x_2	$\cdots$	x_n
度数	f_1	f_2	$\cdots$	f_n

(x_1 が f_1 個，$\cdots$，x_n が f_n 個ある)

この場合に全個数 N は

$$N = f_1 + f_2 + \cdots + f_n \text{ (全個数)}$$

であるから平均 m は

$$m = \frac{1}{N}\{x_1\times f_1 + x_2\times f_2 + \cdots + x_n\times f_n\} \text{ (平均)}$$

となる。

問 この場合の分散 v と標準偏差 σ を求める式を書け。

< 資料の標準化 >

n 個の資料 $\{x_1, x_2, \cdots, x_n\} = \{x_k : k = 1, 2, \cdots, n\}$ の平均を m，標準偏差を σ とする。
このとき

$$\boxed{x_k^* = \frac{x_k - m}{\sigma}} \quad (k = 1, 2, \cdots, n)$$

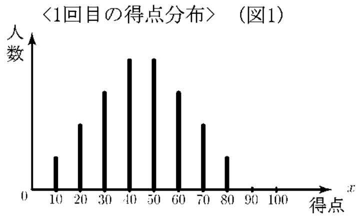
を**標準化した値**という。

例　前ページの例を考える。

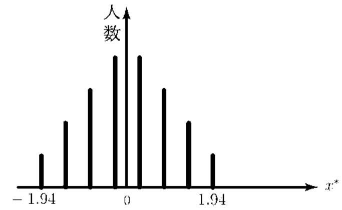
(1)　1 回目の平均は m =45(点)，標準偏差は σ =18.03
80 点を標準化すると
$x = 80 \Rightarrow x^* = \frac{x-m}{\sigma} = \frac{80-45}{18.03} \fallingdotseq 1.94$
以下同様に計算したものを表にする。

得点 x	10	20	30	40	50	60	70	80
標準化した値 x^*	−1.94	−1.39	−0.83	−0.28	0.28	0.83	1.39	1.94
人数	2	4	6	8	8	6	4	2

〈1回目の得点分布〉 (図1)



〈標準化した得点の分布〉 (図2)

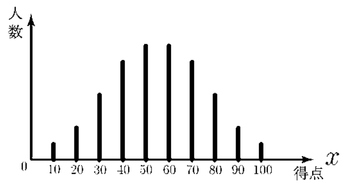


この標準化した値の分布を図 2 に書いた。−1.94 から 1.94 まで
8 個の値が等間隔に並んでいる。これは元の得点分布が 10 点
おきに等間隔に並んでいるからである。

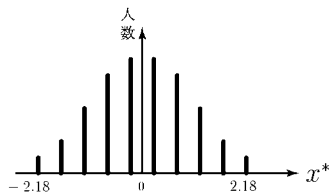
(2)　2 回目の平均は m =55(点)，標準偏差は σ =20.62
90 点を標準化すると
$x = 90 \Rightarrow x^* = \frac{x-m}{\sigma} = \frac{90-55}{20.62} \fallingdotseq 2.18$
以下同様に計算したものを表にする。

得点 x	10	20	30	40	50	60	70	80	90	100
標準化 x^*	−2.18	−1.70	−1.21	−0.73	−0.24	0.24	0.73	1.21	1.70	2.18
人数	1	2	4	6	7	7	6	4	2	1

〈2回目の得点分布〉 (図3)



〈標準化した得点の分布〉 (図4)



この標準化した値の分布を図 4 に書いた。図 4 の分布は
後で述べる標準正規分布 (36 ページ) に似ている。

問　1 回目のテストの標準化した値 x^* が −1 から 1 の範囲にある場合は，元の得点 x が
30 点から 60 点の範囲であるから，その範囲の人数は $6 + 8 + 8 + 6 = 28$(人) であり
全体の 40 分の $28 = \frac{28}{40} = 0.7 = 70(\%)$ である。つまり

$$\boxed{\text{1 回目のテストで } -1 \leqq x^* \leqq 1 \text{ の範囲にある割合は全体の } 70\ \%}$$

といえる。2 回目のテストで以下の範囲にある割合を％で表せ。

(1) (2 回目)$-1 \leqq x^* \leqq 1$ の範囲の割合　　(2) (2 回目)$-2 \leqq x^* \leqq 2$ の範囲の割合

< 偏差値 >

予備校などの模試を受けると，テストの得点およびその偏差値がわかる。偏差値は次の式で求められる。

$$偏差値 = \frac{得点 - 平均点}{標準偏差} \times 10 + 50$$

前ページの標準化した値 $x^* = \dfrac{x-m}{\sigma}$ を使って書くと

$$偏差値 = x^* \times 10 + 50$$

となる。

例 前ページの例を考える。

(1) 1 回目のテストの偏差値を求める。

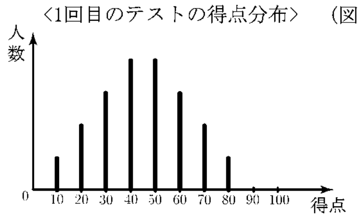
80 点を標準化した値は $x^* = 1.94$ なので

80 点の偏差値 $= 1.94 \times 10 + 50 = 69.4$(点)

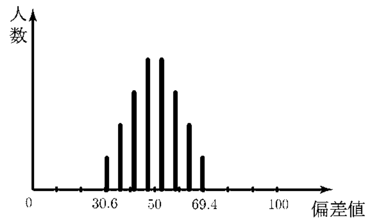
以下同様に計算したものを表にする。

得点 x	10	20	30	40	50	60	70	80
標準化した値 x^*	−1.94	−1.39	−0.83	−0.28	0.28	0.83	1.39	1.94
偏差値	30.6	36.1	41.7	47.2	52.8	58.3	63.9	69.4
人数	2	4	6	8	8	6	4	2

〈1回目のテストの得点分布〉 (図1)



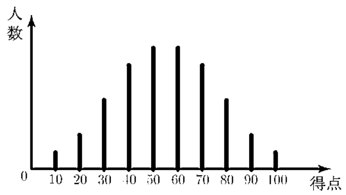
〈1回目のテストの偏差値分布〉 (図2)



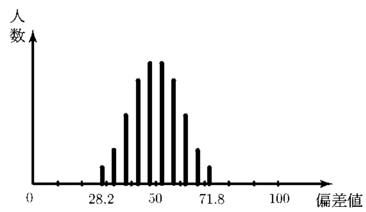
(2) 2 回目のテストの偏差値を表にする。

得点	10	20	30	40	50	60	70	80	90	100
x^*	−2.18	−1.70	−1.21	−0.73	−0.24	0.24	0.73	1.21	1.70	2.18
偏差値	28.2	33	37.9	42.7	47.6	52.4	57.3	62.1	67	71.8
人数	1	2	4	6	7	7	6	4	2	1

〈2回目のテストの得点分布〉 (図3)



〈2回目のテストの偏差値分布〉 (図4)



問 1 例のテストを受けた 3 人の得点が

A 君：1 回目 70 点，2 回目 50 点

B 君：1 回目 50 点，2 回目 70 点

C 君：1 回目 60 点，2 回目 60 点

であった。1 回目と 2 回目の偏差値の平均を A 君，B 君，C 君一人ずつに対し計算せよ。

問 2 平均 45(点)，標準偏差 20 であるテストで，以下の得点の偏差値を求めよ。

(1) 100 点 (2) 60 点 (3) 40 点 (4) 0 点

< 試行と事象 >

世の中には全く偶然によって起こったと思われる現象がたくさんある。これらの現象の一回一回の結果は偶然に支配されていても，数多くの観察するとある法則性が認められる。本書ではこの「法則」を理解することを目的とする。従って本書では特に「数多く観察できる現象」すなわち「再現できる現象」を考察する。

「試行」と「事象」

何回もくり返すことができて，その結果が偶然に支配されているような実験や観察を**試行** (Trail) といい，試行の結果として起こることがらを**事象** (Event) という。

例 1 「1 個のサイコロを投げること」は試行であり,「1 の目が出ること」「偶数の目が出ること」などは事象である。

上の例で「偶数の目が出ること」という事象は「2 の目が出る」,「4 の目が出る」,「6 の目が出る」の 3 つの事象に分けることができる。
しかし「2 の目が出る」という事象は，もうこれ以上分けることができない。

このようにそれ以上分けることのできない事象を**根元事象**という。

上の例で「2 の目が出る」ことを 2 で表すことにすると,「偶数の目が出る」という事象を A とすれば,

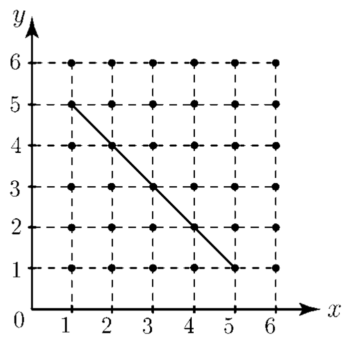
$$A = \{2,\ 4,\ 6\}$$

という集合で表される。またサイコロ投げで根元事象の全体は

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

という集合で表される。このように根元事象の全体からなる事象 U を**全事象**と呼ぶ。

例 2 サイコロを 2 回投げる。1 回目の出た目の数を x 座標 2 回目の出た目の数を y 座標とすると，根元事象は右図のような平面上の点として表される。このとき全事象 U は 36 個の根元事象からなる。集合 U の個数を $n(U)$ で表すと



$$n(U) = 36$$

である。「サイコロを 2 回投げて出た目の和が 6」という事象を A とすれば

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

と表されるから，これは 5 個の根元事象からなる。すなわち $n(A) = 5$ である。

問 サイコロを 3 回投げる。全事象を U とする。「出た目の和が 5」という事象を A とする。$n(U)$ および $n(A)$ を求めよ。

< 有限事象の確率の定義 >

例 1 コインを投げる。このとき“ 表が出る ”か，“ 裏が出る ”か？ 本書では「“ どちらがでやすいか ”という判断ができない」と仮定する。つまり「“ 表が出る可能性 ”と“ 裏が出る可能性 ”が等しい」と仮定する。このようなとき「“ 表が出ること ”と“ 裏が出ること ”は“ **同様に確からしい** ”」という。

例 2 サイコロを投げる。本書では「“ どの目がでやすいか ”という判断ができない」と仮定する。つまり「“ どの目が出るか ”は偶然に支配され，しかも“ それぞれの目の出る可能性は等しい ”」と仮定する。このようなとき「各根元事象 $\{1\}$,$\{2\}$,$\{3\}$,$\{4\}$,$\{5\}$,$\{6\}$ は“ **同様に確からしい** ”」という。

全事象 U が有限個の根元事象の集合で表されるとき，U を**有限事象**という。有限事象の場合に本書では「“ どの根元事象が起こりやすいか ”という判断ができない」と仮定する。つまり「“ どの根元事象が起こるか ”は全くの偶然に支配され，しかも“ 各根元事象の起こる可能性は等しい ”」と仮定する。このとき「各根元事象は“ **同様に確からしい** ”」という。

有限事象の場合に，全事象 U の個数を $n(U)$，事象 A の個数を $n(A)$ で表すとき，

$$\boxed{P(A)=\frac{n(A)}{n(U)}} \qquad \textbf{(事象 } A \textbf{ の起こる確率)}$$

を「事象 A の起こる確率」という。

例題 サイコロを 2 回投げて，出た目の和が 6 になる確率を求めよ。

誤答例 目の和は

$$2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12$$

の 11 通りであるから $\underline{(答)\ \frac{1}{11}}$

正解 サイコロを 2 回投げるとき全事象 U は (前ページ例 2 から)36 個の根元事象からなる。すなわち $n(U)$=36 である。このうち「出た目の和が 6 になる事象」を A とすると

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

の 5 個の根元事象からなる。すなわち $n(A)$=5 である。よって

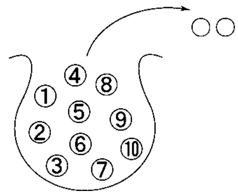
$$\underline{(正答)\ \ P(A)=\frac{n(A)}{n(U)}=\frac{5}{36}}$$

(注) 誤答例の場合“ 目の和 ”は根元事象でない。たとえば「目の和が 3」という事象は「1 回目 1，2 回目 2」と「1 回目 2，2 回目 1」の 2 つの事象に分けられる。「目の和が 6」は 5 個の事象に分けられる。従って，“ 目の和 ”は同様に確からしくない。

問 1 サイコロを 2 回投げて出た目の和が 7 になる確率を求めよ。

問 2 サイコロを 3 回投げて出た目の和が 5 になる確率を求めよ。

＜基本的な確率の計算＞



例題　袋の中に 10 個の玉がはいっていて，1 から 10 まで番号が書いてある。この袋から 2 個とり出すとき 2 個とも (玉の番号が)4 以下である確率を以下のそれぞれの場合について求めよ。

(1) 1 個の玉をとり出し，それを元にもどしてから 2 個目をとり出す場合

(2) 1 個の玉をとり出し，それを元にもどさずに 2 個目をとり出す場合

(3) 同時に 2 個とり出す場合

(解)

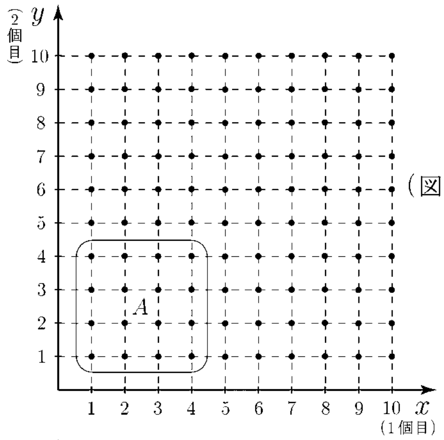
(1) 1 個目の玉の番号を x 座標，2 個目の玉の番号を y 座標で表すと根元事象は図 1 のような平面上の点で表される。この場合全事象 U は 10^2 個の根元事象からなる。つまり

$$n(U) = 10^2 = 100$$

である。2 個とも 4 以下の場合を事象 A とすると

$$n(A) = 4^2 = 16$$

より $P(A) = \dfrac{16}{100} = \dfrac{4}{25}$　　(答) $\dfrac{4}{25}$



(図1)

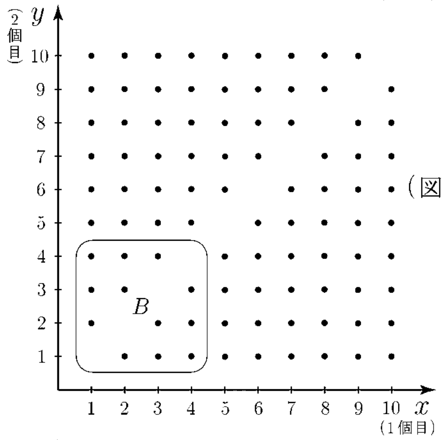
(2) 1 個目を元にもどさずに 2 個目をとり出す場合，根元事象の全体は図 2 のような点で表される。このとき全事象 U の個数は，10 個から 2 個順にとり出す順列の数であるから

$$n(U) = {}_{10}P_2 = 10 \times 9 = 90$$

2 個とも 4 以下の事象 B の個数は，4 個から 2 個順にとり出す順列の数だから

$$n(B) = {}_4P_2 = 4 \times 3 = 12$$

より $P(B) = \dfrac{12}{90} = \dfrac{2}{15}$　　(答) $\dfrac{2}{15}$



(図2)

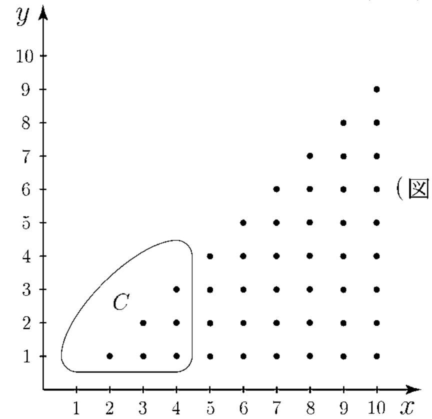
(3) 同時に 2 個とり出す場合，全事象 U の個数は 10 個から 2 個とり出す組合せの数だから

$$n(U) = {}_{10}C_2 = \frac{10 \times 9}{2 \times 1} = 45$$

2 個とも 4 以下の事象 C の個数は，4 個から 2 個とり出す組合せの数だから

$$n(C) = {}_4C_2 = \frac{4 \times 3}{2 \times 1} = 6$$

より $P(C) = \dfrac{6}{45} = \dfrac{2}{15}$　　(答) $\dfrac{2}{15}$



(図3)

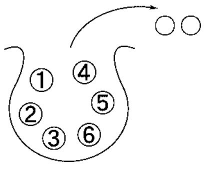
問　例の場合に袋から 3 個とり出し，3 個とも 4 以下の番号である確率を以下の各場合について求めよ。

(1) とり出した玉をそのつど元にもどす場合　(2) とり出した玉を元にもどさない場合　(3) 同時に 3 個とり出す場合

< 共通事象の確率 >

全事象 U に対し事象 A や事象 B は U の部分集合と考えられる。
A と B の共通部分集合 $A \cap B$ は「**A と B がともに起こる事象**」を表す。
これを「**A と B の共通事象**」という。

例 1 袋の中に 6 個の玉がはいっていて，1 から 6 まで番号が書いてある。
この袋から順に 1 個づつ玉をとりだす。事象 A と B を



事象 A : 1 個目 4 以下の番号の玉が出る事象

事象 B : 2 個目 5 以上の番号の玉が出る事象

とするとき，A と B の共通事象の起こる確率 $P(A \cap B)$ を以下の各場合について考える。

(1) 「1 個の玉をとり出し，それを元にもどしてから 2 個目をとり出す場合」
前ページと同様に全事象 U を図 1 のように表すと，

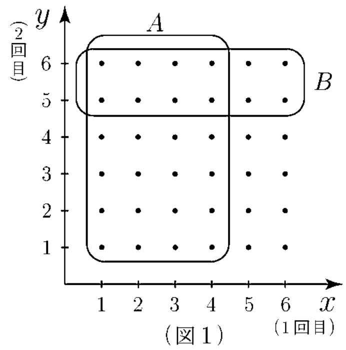
$$U = \{(x, y) : x \text{ と } y \text{ は } 1 \text{ から } 6 \text{ までの整数}\}$$

と書ける。従って $n(U) = 6^2 = 36$ である。A と B の共通事象は

$$A \cap B = \{(1,5), (1,6), (2,5), (2,6), (3,5), (3,6), (4,5), (4,6)\}$$

であるから $n(A \cap B) = 4 \times 2 = 8$ である。よって，

$$P(A \cap B) = \frac{n(A \cap B)}{n(U)} = \frac{8}{36} = \frac{2}{9}$$



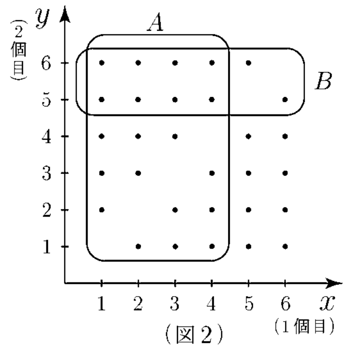
（図 1）

(2) 「1 個の玉をとり出し，それを元にもどさずに 2 個目をとり出す場合」
前ページ (2) の場合と同様に全事象 U を図 2 のように表すと

$$n(U) = {}_6P_2 = 6 \times 5 = 30$$

である。一方共通事象 $A \cap B$ は (1) と同じであるから $n(A \cap B) = 8$
より

$$P(A \cap B) = \frac{n(A \cap B)}{n(U)} = \frac{8}{30} = \frac{4}{15}$$



（図 2）

例 2 サイコロを 3 回ふる。事象 A, B, C が以下の場合に共通事象 $A \cap B \cap C$ の確率を求める。

事象 A : 1 回目 2 以下の目が出る事象

事象 B : 2 回目 3 または 4 の目が出る事象

事象 C : 3 回目 5 以上の目が出る事象

全事象 U は例 1(1) と同様に

$$U = \{(x, y, z) : x, y, z \text{ は } 1 \text{ から } 6 \text{ までの整数}\}$$

と表される。従って $n(U) = 6^3 = 216$ である。このとき $A \cap B \cap C$ は

$$A \cap B \cap C = \{(1,3,5), (1,3,6), (1,4,5), (1,4,6), (2,3,5), (2,3,6), (2,4,5), (2,4,6)\}$$

となるので $n(A \cap B \cap C) = 8$ より $P(A \cap B \cap C) = \dfrac{8}{216} = \dfrac{1}{27}$

問 例 2 の場合に事象 A' を

A' : 1 回目 4 以下の目が出る事象

とするとき，$P(A' \cap B \cap C)$ を求めよ。

< 独立試行 1 >

例 1　11 ページ例題の (1) や 12 ページ例 1(1) のようにとり出した玉をそのつど元にもどす抽出方法を**復元抽出**といい，とり出した玉を元にもどさない抽出方法を**非復元抽出**という。非復元抽出では 1 回目の抽出結果が 2 回目の抽出に影響を与えるが，復元抽出の場合は 1 回目の抽出結果が 2 回目に影響することはない。

「独立な試行」

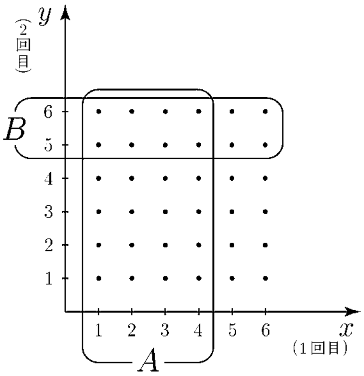
いくつかの試行において，それらの結果が互いに影響しないとき，これらの試行は**「互いに独立」**であるという。

例 2　(1)　復元抽出の場合，各回の試行 $\big(=$ 抽出$\big)$ は互いに独立である。
(2)　サイコロを 3 回振ったとき，各回の試行はお互いに独立である。

例 3　サイコロを 2 回振る。以下の事象 A , B について共通事象 $A\cap B$ の確率を考える。

事象 A : 1 回目 4 以下の目が出る事象

事象 B : 2 回目 5 以上の目が出る事象



前ページ例 1(1) と同様にして

「A と B が共に起こる確率」 $=\ P(A\cap B)=\dfrac{n(A\cap B)}{n(U)}=\dfrac{4\times 2}{6\times 6}=\dfrac{2}{9}$

である。一方

$$P(A)=\frac{4\times 6}{6\times 6}=\frac{2}{3}\quad ,\quad P(B)=\frac{6\times 2}{6\times 6}=\frac{1}{3}$$

であるから

$$P(A\cap B)=\frac{2}{9}=\frac{2}{3}\times\frac{1}{3}=P(A)\times P(B)$$

が成り立つ。

一般に，独立な試行の場合，次の**積の法則**がなりたつ。

2 つの試行が互いに独立な試行で，それぞれの事象が A , B のとき $P(A\cap B)=P(A)\times P(B)$

(積の法則)

問 1　前ページ例 1 の (2)(非復元抽出) の場合に $P(A\cap B)$ と $P(A)\times P(B)$ を求めよ。

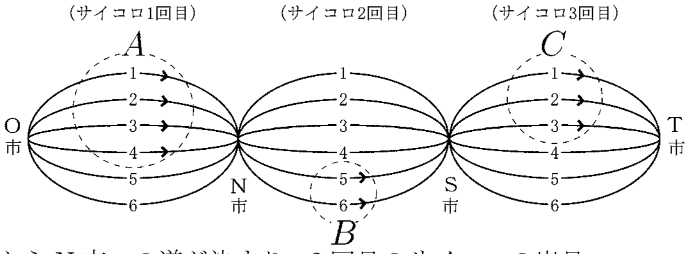
問 2　サイコロを 2 回振る。以下の事象 A , B に対し $P(A\cap B)$ と $P(A)\times P(B)$ を求めよ。

A : 1 回目 3 以下の目が出る事象

B : 2 回目 5 以上の目が出る事象

< 独立試行 2 >

例 1 O 市から出発し，N 市，S 市の順で T 市へ行く，道は右図のように各 6 通りあり，サイコロを振って通る道を決める。



1 回目のサイコロの出た目の数で O 市から N 市への道が決まり，2 回目のサイコロの出目で N 市から S 市への道が決まり，3 回目のサイコロの出目で S 市から T 市への道が決まる。右図のように道の番号があるとき，以下の事象 A,B,C を考える。

事象 A ： O 市から N 市へ行くとき 4 以下の番号の道を通る事象
事象 B ： N 市から S 市へ行くとき 5 以上の番号の道を通る事象
事象 C ： S 市から T 市へ行くとき 3 以下の番号の道を通る事象

このとき O 市から T 市へ行く全ての道順を全事象 U とすれば,

$$n(U) = 6^3 = 216\ ,\quad n(A \cap B \cap C) = 4 \times 2 \times 3 = 24$$

であるから

$$P(A \cap B \cap C) = \frac{n(A \cap B \cap C)}{n(U)} = \frac{24}{216} = \frac{1}{9}$$

となる。一方

$$P(A) = \frac{n(A)}{n(U)} = \frac{4 \times 6 \times 6}{6^3} = \frac{2}{3}\ ,\quad P(B) = \frac{6 \times 2 \times 6}{6^3} = \frac{1}{3}\ ,\quad P(C) = \frac{6 \times 6 \times 3}{6^3} = \frac{1}{2}$$

だから

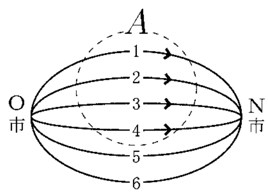
$$P(A \cap B \cap C) = \frac{1}{9} = \frac{2}{3} \times \frac{1}{3} \times \frac{1}{2} = P(A) \times P(B) \times P(C)$$

がなりたつ。

一般に独立試行の場合，**積の法則**は 3 回以上の試行の場合にも適用される。

> n 回の試行が互いに独立な試行で，それぞれの事象が $A_1, A_2, \cdots, A_n$ のとき
> $$P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_1) \times P(A_2) \times \cdots \times P(A_n)$$

(注) 例 1 の場合に $P(A)$ を求めるとき，O 市から N 市への道順だけを考えて，$P(A) = \frac{4}{6} = \frac{2}{3}$ と計算してよい。
独立試行の場合は，各回の確率を独立に計算してよい。



例 2 サイコロを 3 回振る。事象 A, B, C を

A ： 1 回目 4 以上の目が出る事象
B ： 2 回目 2 以下の目が出る事象
C ： 3 回目奇数の目が出る事象

とすると

$P(A) = \frac{3}{6} = \frac{1}{2}$， $P(B) = \frac{2}{6} = \frac{1}{3}$， $P(C) = \frac{3}{6} = \frac{1}{2}$

より

$P(A \cap B \cap C) = P(A) \times P(B) \times P(C) = \frac{1}{2} \times \frac{1}{3} \times \frac{1}{2} = \frac{1}{12}$

問 サイコロを 4 回振る。事象 A, B, C, D を

A ： 1 回目 4 以下の目が出る事象
B ： 2 回目偶数の目が出る事象
C ： 3 回目 3 の倍数の目が出る事象
D ： 4 回目 5 が出る事象

とするとき $P(A \cap B \cap C \cap D)$ を求めよ。

< 反復試行 >

サイコロを何回も振るとか，袋から玉の復元抽出 (取り出しては元に戻す) を繰り返すというように，同じ実験や観察を繰り返し行なうことを **反復試行** という。

例 1 サイコロを 2 回振って 1 の目が 1 回出る確率を求めたい。

この場合は次の 2 通りの事象に分けられる。

事象 A ： 1 回目 1 の目が出て，2 回目 1 以外の目が出る事象

事象 B ： 1 回目 1 以外の目が出て，2 回目 1 の目が出る事象

1 回目と 2 回目の試行は独立だから，

$P(A) =$「1 回目 1 の目が出る確率」×「2 回目 1 以外の目が出る確率」$= \frac{1}{6} \times \frac{5}{6} = \frac{5}{36}$

であり，同様にして $P(B) = \frac{5}{36}$ である。求める確率は

「2 回のうち 1 の目が 1 回出る確率」$= P(A) + P(B) = \frac{5}{36} + \frac{5}{36} = \frac{5}{18}$

(注) このとき $A \cap B = \phi$ となるので $P(A \cup B) = P(A) + P(B)$ になる。このような事象 A, B を互いに**排反**という。

例 2 サイコロを 4 回振って 1 の目が 2 回出る確率を求めたい。

この場合は右の表のように 6 通りの場合がある。

事象 A_1 を「1 回目・2 回目ともに 1 の目が出て，かつ 3 回目・4 回目ともに 1 以外の目が出る事象」とすると，各回は独立だから，A_1 の起こる確率は

$$P(A_1) = \frac{1}{6} \times \frac{1}{6} \times \frac{5}{6} \times \frac{5}{6} = \left(\frac{1}{6}\right)^2 \times \left(\frac{5}{6}\right)^2$$

となる。以下同様にして $P(A_2) = \cdots = P(A_6) = \left(\frac{1}{6}\right)^2 \times \left(\frac{5}{6}\right)^2$

であるから求める確率は

$$P(A_1) + \cdots + P(A_6) = 6 \times \left(\frac{1}{6}\right)^2 \times \left(\frac{5}{6}\right)^2 = \frac{25}{216}$$

事象＼回数	1	2	3	4	確率
A_1	○	○	×	×	$\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^2$
A_2	○	×	○	×	$\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^2$
A_3	○	×	×	○	$\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^2$
A_4	×	○	○	×	$\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^2$
A_5	×	○	×	○	$\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^2$
A_6	×	×	○	○	$\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^2$

(1の目が出るとき○、1の目が出ないとき×)

(注) 例 2 の場合わけの数が “ 6 通り ” なのは以下の理由による，1 回目から 4 回目のうち，1 の目の出る 2 回を決める場合の数，つまり 1，2，3，4 から 2 個取った組合せの数は ${}_4C_2 = \frac{4\times3}{2\times1} = 6$(通り) である。

例 3 サイコロを 5 回振って 1 の目が 2 回出る確率は ${}_5C_2 \times \left(\frac{1}{6}\right)^2 \times \left(\frac{5}{6}\right)^3 = \frac{625}{3888}$ である。

例 4 サイコロを 5 回振って 3 の倍数が 4 回出る確率は ${}_5C_4 \times \left(\frac{1}{3}\right)^4 \times \left(\frac{2}{3}\right)^1 = \frac{5\times2}{3^5} = \frac{10}{243}$ である。

問 次の確率を求めよ。

(1) サイコロを 5 回振って 1 の目が 3 回出る確率

(2) コインを 5 回投げて表が 2 回出る確率

＜確率変数＞

例 1　サイコロを 4 回振って 1 の目の出る回数を X とする。このとき

「サイコロを 4 回振って 1 の目が 2 回出る確率」

を単に

$$P(X=2)$$

と略記する。前ページ例 2 より

$$P(X=2) = {}_4C_2 \times \left(\frac{1}{6}\right)^2 \times \left(\frac{5}{6}\right)^2 = \frac{150}{1296} \left(= \frac{25}{216}\right)$$

となる。$P(X=0)$ は「4 回とも 1 の目が出ない確率」だから

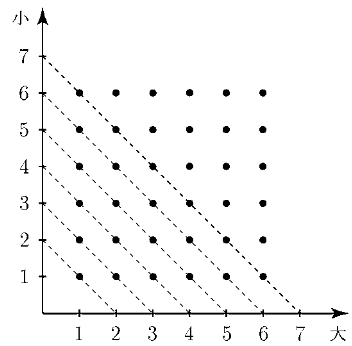
$$P(X=0) = {}_4C_0 \times \left(\frac{5}{6}\right)^4 = \frac{625}{1296}$$

以下同様に計算したものを右の表にした。

(表 1)

1の目の回数 X	0	1	2	3	4	計
確率	$\frac{625}{1296}$	$\frac{500}{1296}$	$\frac{150}{1296}$	$\frac{20}{1296}$	$\frac{1}{1296}$	1

例 2　大小 2 つのサイコロを振る。X を目の和とすると，X は 2 から 12 までの値をとり得る。各値に対する確率を表 2 にした。



(表 2)

目の和 X	2	3	4	5	6	7	8	9	10	11	12	計
確率	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	1

例 1 や例 2 のような X を**確率変数** (random variable) といい，表 1 や表 2 を**確率分布**という。

例 3　例 2 の場合を考える。

「目の和が 6 以上 8 以下である確率」

を単に

$$P(6 \leqq X \leqq 8)$$

と略記する。表 2 より

$$P(6 \leqq X \leqq 8) = P(X=6) + P(X=7) + P(X=8) = \frac{5}{36} + \frac{6}{36} + \frac{5}{36} = \frac{16}{36} = \frac{4}{9}$$

問　コインを 5 回投げた。表の出る回数を X とする。前ページの結果より

$$P(X=2) = {}_5C_2 \times \left(\frac{1}{2}\right)^2 \times \left(\frac{1}{2}\right)^3 = \frac{10}{32}$$

となる。右の確率分布表を完成し，$P(2 \leqq X \leqq 4)$ を求めよ。

表の回数 X	0	1	2	3	4	5	計
確 率			$\frac{10}{32}$				1

＜ 期待値 ＞

例 1　表 1 のような 100 本の宝くじの賞金総額は

(100本のくじ)

	1等	2等	3等
賞金(円)	5000	500	100
本数	2	8	90

(表1)

$$5000 \times 2 + 500 \times 8 + 100 \times 90 = 23000 \text{ (円)}$$

である。宝くじ 1 本あたりの賞金額の平均値は

$$\frac{5000 \times 2 + 500 \times 8 + 100 \times 90}{100} = 230$$

よって，平均値は 230 円である。くじ 1 枚につき平均 230 円期待できる。

この値はくじ 1 枚の値段を決めるときなどに利用される。このような値のことを一般に**期待値** (Expectation) という。この値が金額のときは**期待金額**ともいう。

例 2　例 1 の賞金を X とすると，表 2 のような確率分布になる。上の期待値を求める計算は

賞金 X	5000	500	100	計
当たる確率	$\frac{2}{100}$	$\frac{8}{100}$	$\frac{90}{100}$	1

(表2)

$$\text{期待値} = 5000 \times \frac{2}{100} + 500 \times \frac{8}{100} + 100 \times \frac{90}{100} = 230$$

と書ける。つまり（賞金）×（当たる確率）の和になっている。

一般に確率変数 X の確率分布が表 3 のような場合に，X の期待値 $E[X]$ を次式で定める。

X	x_1	x_2	$\cdots$	x_n	計
確率	p_1	p_2	$\cdots$	p_n	1

(表3)

$$\boxed{E[X] = x_1p_1 + x_2p_2 + \cdots + x_np_n} \quad (X \text{ の期待値})$$

例 3　100 円硬貨 4 枚を投げて表の出た硬貨の金額を X(円) とする。X の確率分布は表 4 のようになるから，期待値 $E[X]$ は

X	0	100	200	300	400	計
確率 p	$\frac{1}{16}$	$\frac{4}{16}$	$\frac{6}{16}$	$\frac{4}{16}$	$\frac{1}{16}$	1

(表4)

$$E[X] = 0 \times \frac{1}{16} + 100 \times \frac{4}{16} + 200 \times \frac{6}{16} + 300 \times \frac{4}{16} + 400 \times \frac{1}{16} = 200 \text{ (円)}$$

問　100 円硬貨 5 枚を投げて表の出た硬貨の金額を X(円) とする。前ページ問の結果を用いて右に確率分布表を完成し，期待値 $E[X]$ を求めよ。

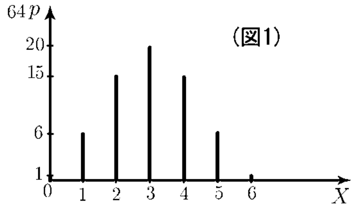
X	0	100	200	300	400	500	計
p							1

< 確率変数の平均と分散 1 >

例 コインを 6 回投げて表の出た回数を X とする。X の確率分布 (表 1) に対し，確率を 64 倍にしたものを棒グラフにしたのが図 1 である。

(表 1)

X	0	1	2	3	4	5	6	計
確率 p	$\frac{1}{64}$	$\frac{6}{64}$	$\frac{15}{64}$	$\frac{20}{64}$	$\frac{15}{64}$	$\frac{6}{64}$	$\frac{1}{64}$	1
$64p$	1	6	15	20	15	6	1	64



(図1)

6 ページと同様にして，図 1 の棒グラフの平均 m と分散 v を求める。

$$m = \frac{1}{64}\{0\times1 + 1\times6 + 2\times15 + 3\times20 + 4\times15 + 5\times6 + 6\times1\} = 3$$

より平均は $m = 3$ である。一方表 1 から X の期待値は

$$E[X] = 0\times\frac{1}{64} + 1\times\frac{6}{64} + 2\times\frac{15}{64} + 3\times\frac{20}{64} + 4\times\frac{15}{64} + 5\times\frac{6}{64} + 6\times\frac{1}{64} = 3$$

より期待値は $E[X] = 3$ である。すなわち，平均と期待値は一致する。そこで X の期待値 $E[X]$ を確率変数 X の**平均**ともいう。

一方分散 v も 6 ページと同様に計算すると

$$v = \frac{1}{64}\{(0-3)^2\times1 + (1-3)^2\times6 + (2-3)^2\times15 + (3-3)^2\times20 + (4-3)^2\times15 + (5-3)^2\times6 + (6-3)^2\times1\} = \frac{3}{2}$$

より分散は $v = \frac{3}{2}$ である。この値は図 1 のデータの広がりぐあいを表す値である。そこでこの値 $v = \frac{3}{2}$ を確率変数 X の**分散**と定める。

一般に確率変数 X の確率分布が表 2 の場合に X の平均 (= 期待値)m と分散 v を次の式で定める。

(表 2)

X	x_1	x_2	$\cdots$	x_n	計
確率 p	p_1	p_2	$\cdots$	p_n	1

$$\boxed{m = E[X] = x_1p_1 + x_2p_2 + \cdots + x_np_n}$$ **平均 (期待値)**

$$\boxed{v = (x_1-m)^2p_1 + (x_2-m)^2p_2 + \cdots + (x_n-m)^2p_n}$$ **分散**

(注) 例の場合に分散 v を求める式は

$$v = (0-3)^2\times\frac{1}{64} + (1-3)^2\times\frac{6}{64} + (2-3)^2\times\frac{15}{64} + (3-3)^2\times\frac{20}{64} + (4-3)^2\times\frac{15}{64} + (5-3)^2\times\frac{6}{64} + (6-3)^2\times\frac{1}{64}$$

と書き直す事が出来る。

問 コインを 4 回投げて表の出た回数を X とする。右の確率分布表を完成し，平均 m と分散 v を求めよ。

X	0	1	2	3	4	計
確率 p						1

< 確率変数の平均と分散 2 >

X	1	2	3	計
確率 p	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{6}$	1

(表 1)

例 1 表 1 のような確率分布をした確率変数 X の期待値は

$$E[X] = 1 \times \frac{1}{2} + 2 \times \frac{1}{3} + 3 \times \frac{1}{6} = \frac{5}{3}$$

である。X に対し

$$Y = 6X + 4$$

とおくと Y も確率変数でありその期待値は表 2 より

X	1	2	3	計
$Y = 6X+4$	10	16	22	
確率 p	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{6}$	1

(表 2)

$$E[Y] = 10 \times \frac{1}{2} + 16 \times \frac{1}{3} + 22 \times \frac{1}{6} = 14$$

となる。一方 $6E[X] + 4 = 6 \times \frac{5}{3} + 4 = 14$ より

$$E[Y] = E[6X+4] = 6E[X] + 4$$

が成り立つ。

一般に確率変数 X と定数 a, b に対し

$$\boxed{E[aX+b] = aE[X] + b} \quad (a \text{ と } b \text{ は定数})$$

が成り立つ。

確率分布が表 3 である確率変数 X の平均を m，分散を v とする。すなわち

X	x_1	x_2	$\cdots$	x_n	計
確率 p	p_1	p_2	$\cdots$	p_n	1

(表 3)

$$m = E[X] = x_1p_1 + x_2p_2 + \cdots + x_np_n$$

$$v = (x_1 - m)^2 p_1 + (x_2 - m)^2 + \cdots + (x_n - m)^2 p_n$$

である。X に対し

$$(X - m)^2$$

$(X-m)^2$	$(x_1-m)^2$	$(x_2-m)^2$	$\cdots$	$(x_n-m)^2$	計
確率 p	p_1	p_2	$\cdots$	p_n	1

(表 4)

も確率変数であり，その確率分布は表 4 になる。これから $(X-m)^2$ の期待値が v

$$\boxed{E\big[(X-m)^2\big] = (x_1-m)^2p_1 + (x_2-m)^2p_2 + \cdots + (x_n-m)^2p_n = v} \quad (\text{「 } \boldsymbol{X} \textbf{の分散」} = \textbf{「}(\boldsymbol{X}-\boldsymbol{m})^2 \textbf{の平均」})$$

であることがわかる。X の分散を記号 $V(X)$ で書くと

$$\boxed{V(X) = E\big[(X-m)^2\big] = E\Big[\big(X - E(X)\big)^2\Big]} \quad (\boldsymbol{X}\textbf{の分散})$$

となる。ここで $m = E[X]$ である。

例 2 平均 m 分散 v の確率変数 X に対し $Y = 6X + 4$ とおくと
$E[Y] = E[6X+4] = 6E[X] + 4 = 6m + 4$ より Y の平均は $6m + 4$ である。Y の分散は

$$V(Y) = E\Big[\big(Y - E[Y]\big)^2\Big] = E\Big[\big(6X + 4 - (6m+4)\big)^2\Big] = E\big[6^2(X-m)^2\big] = 6^2E\big[(X-m)^2\big] = 36v$$

問 平均 m，分散 v の確率変数 X と定数 a，b に対し，$Y = aX + b$ とおく。
$E[Y]$ と $V(Y)$ を求めよ。

＜独立確率変数 1＞

例 1 サイコロを 2 回投げる。確率変数 X と Y を

X： 1 回目の出た目の数
Y： 2 回目の出た目の数

とする。今「1 回目に 3 の目が出てかつ 2 回目に 4 の目が出る確率」を単に

$$P(X=3,\ Y=4)$$

と書くことにすると

$$P(X=3,\ Y=4)=\frac{1}{36}=\frac{1}{6}\times\frac{1}{6}=P(X=3)\times P(Y=4)$$

となる。以下同様にして全ての $1 \leqq i, j \leqq 6$ に対し

$$P(X=i,\ Y=j)=P(X=i)\times P(Y=j) \qquad (1\leqq i\leqq 6,\ 1\leqq j\leqq 6)$$

がなりたつ。

一般に 2 つの確率変数 X と Y に対し

$$\boxed{P(X=x,\ Y=y)=P(X=x)\times P(Y=y)}$$

が全ての x,y に対し成り立つとき,「X と Y は **独立**」という。

例 1 の場合 X は 1 回目の試行によって定まり，Y は 2 回目の試行によって定まる。1 回目の試行と 2 回目の試行が独立だから，X と Y は独立になる。

このように 独立な試行によって定まる確率変数は独立になる。

例 2 10 円硬貨 2 枚と 100 円硬貨 1 枚を同時に投げる。

X ： 10 円硬貨の表の枚数
Y ： 100 円硬貨の表の枚数

X	0	1	2	計
確率	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$	1

(表 1)

Y	0	1	計
確率	$\frac{1}{2}$	$\frac{1}{2}$	1

(表 2)

10 円硬貨を投げることと 100 円硬貨を投げることは互いに独立な試行である。従って X と Y は独立。よって

$$P(X=0,\ Y=1)=P(X=0)\times P(Y=1)=\frac{1}{4}\times\frac{1}{2}=\frac{1}{8}$$

$$P(X=1,\ Y=0)=P(X=1)\times P(Y=0)=\frac{1}{2}\times\frac{1}{2}=\frac{1}{4}$$

問 1 例 2 の結果を表 3 に記した。表 3 を X と Y の**同時確率分布**という。表 3 を完成せよ。

X	0	0	1	1	2	2	計
Y	0	1	0	1	0	1	
確率		$\frac{1}{8}$	$\frac{1}{4}$				1

(表 3)

問 2 例 2 の場合，表 3 より

$$P(X+Y=1)=P(X=0,\ Y=1)+P(X=1,\ Y=0)=\frac{1}{8}+\frac{1}{4}=\frac{3}{8}$$

となる。表 4 を完成せよ。

$X+Y$	0	1	2	3	計
確率		$\frac{3}{8}$			1

(表 4)

＜独立確率変数 2＞

例 1 前ページ例 2 の場合，平均は

$$E[X]=1\ ,\quad E[Y]=\frac{1}{2}\ ,\quad E[X+Y]=\frac{3}{2}$$

$X+Y$	0	1	2	3	計
確率	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$	1

であり分散は

$$V(X)=(0-1)^2\times\frac{1}{4}+(1-1)^2\times\frac{1}{2}+(2-1)^2\times\frac{1}{4}=\frac{1}{2}\ ,\quad V(Y)=\left(0-\frac{1}{2}\right)^2\times\frac{1}{2}+\left(1-\frac{1}{2}\right)^2\times\frac{1}{2}=\frac{1}{4}$$

$$V(X+Y)=\left(0-\frac{3}{2}\right)^2\times\frac{1}{8}+\left(1-\frac{3}{2}\right)^2\times\frac{3}{8}+\left(2-\frac{3}{2}\right)^2\times\frac{3}{8}+\left(3-\frac{3}{2}\right)^2\times\frac{1}{8}=\frac{3}{4}$$

となる。

一般に 2 つの確率変数 X と Y に対して

$$\boxed{E[X+Y]=E[X]+E[Y]}\ \text{(確率変数の和の平均)}$$

が成り立つ。さらに X と Y が独立ならば

$$\boxed{V(X+Y)=V(X)+V(Y)}\ \text{(独立確率変数の和の分散)}$$

が成り立つ。(証明略)

3 個以上の確率変数についても同様な結果が成り立つ。

n 個の独立な試行によって定まる n 個の確率変数 $X_1,\cdots,X_n$ に対し

$$P(X_1=x_1,X_2=x_2,\cdots,X_n=x_n)=P(X_1=x_1)\times P(X_2=x_2)\times\cdots\times P(X_n=x_n)$$

がすべての実数値 $x_1,\cdots,x_n$ に対して成立する。このとき確率変数列 $X_1,\cdots,X_n$ は互いに**独立**という。このとき上と同様に

$$\boxed{\begin{aligned}&E[X_1+X_2+\cdots+X_n]=E[X_1]+E[X_2]+\cdots+E[X_n]\\&V(X_1+X_2+\cdots+X_n)=V(X_1)+V(X_2)+\cdots+V(X_n)\end{aligned}}\quad\left(\begin{array}{l}\text{独立確率変数列の和}\\\text{の平均と分散}\end{array}\right)$$

が成り立つ。

例 2 サイコロを 3 回振る。確率変数 X_1,X_2,X_3 を以下のように定める

X_1 : 1 回目 1 の目が出れば $X_1=1$，そうでなければ $X_1=0$

X_2 : 2 回目 1 の目が出れば $X_2=1$，そうでなければ $X_2=0$

X_3 : 3 回目 1 の目が出れば $X_3=1$，そうでなければ $X_3=0$

X_1,X_2,X_3 は互いに独立である。またその確率分布は全て同じであるので右表のように書いた。

$X_k(k=1,2,3)$ の確率分布

X_k	0	1	計
確率	$\frac{5}{6}$	$\frac{1}{6}$	1

$$E[X_k]=0\times\frac{5}{6}+1\times\frac{1}{6}=\frac{1}{6}\quad ,\quad V(X_k)=\left(0-\frac{1}{6}\right)^2\times\frac{5}{6}+\left(1-\frac{1}{6}\right)^2\times\frac{1}{6}=\frac{5}{36}\qquad(k=1,2,3)$$

より

$$E[X_1+X_2+X_3]=E[X_1]+E[X_2]+E[X_3]=\frac{1}{6}+\frac{1}{6}+\frac{1}{6}=\frac{1}{2}$$

$$V(X_1+X_2+X_3)=V(X_1)+V(X_2)+V(X_3)=\frac{5}{36}+\frac{5}{36}+\frac{5}{36}=\frac{5}{12}$$

問 サイコロを n 回振る。例 2 のように $X_1,X_2,\cdots,X_n$ を定めるとき和の平均と分散を求めよ。

$E[X_1+X_2+\cdots+X_n]=$ 　　　　　　　　, 　$V(X_1+X_2+\cdots+X_n)=$

＜ 二項分布 ＞

例 1 サイコロを 4 回振る。1 の目の出る回数を X とすると，15 ページより X の確率分布は右表のようになる。

X	0	1	2	3	4	計
確率	${}_4C_0(\frac{1}{6})^0(\frac{5}{6})^4$	${}_4C_1(\frac{1}{6})^1(\frac{5}{6})^3$	${}_4C_2(\frac{1}{6})^2(\frac{5}{6})^2$	${}_4C_3(\frac{1}{6})^3(\frac{5}{6})^1$	${}_4C_4(\frac{1}{6})^4(\frac{5}{6})^0$	1

前ページ例 2 のように

$$\boxed{X_k : k \text{ 回目に } 1 \text{ の目が出れば } X_k = 1, \text{ そうでなければ } X_k = 0} \quad (k = 1, 2, 3, 4)$$

とおくと，4 回振って 1 の目が出る回数 X は

$$X = X_1 + X_2 + X_3 + X_4$$

と表される。X_1, X_2, X_3, X_4 は互いに独立であり，$E[X_k] = \dfrac{1}{6}$，$V(X_k) = \dfrac{5}{36}$ より

$$E[X] = E[X_1] + E[X_2] + E[X_3] + E[X_4] = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = 4 \times \frac{1}{6} = \frac{2}{3}$$

$$V(X) = V(X_1) + V(X_2) + V(X_3) + V(X_4) = \frac{5}{36} + \frac{5}{36} + \frac{5}{36} + \frac{5}{36} = 4 \times \frac{5}{36} = \frac{5}{9}$$

例 2 サイコロを n 回振る。1 の目の出る回数を X とすると，1 の目が k 回出る確率は

$$P(X = k) = {}_nC_k \left(\frac{1}{6}\right)^k \left(\frac{5}{6}\right)^{n-k}, \quad (k = 0, 1, 2, \cdots, n)$$

となる。例 1 と同様にして，X の平均と分散は

$$E[X] = E[X_1] + \cdots + E[X_n] = n \times \frac{1}{6}$$

$$V(X) = V(X_1) + \cdots + V(X_n) = n \times \frac{5}{36} = n \times \frac{1}{6} \times \left(1 - \frac{1}{6}\right)$$

となる。

一般にある試行において事象 A の起こる確率を p $(0 < p < 1)$ とする。この試行を n 回くり返すとき，事象 A の起こる回数を X とすると，k 回起こる確率は

$$\boxed{P(X = k) = {}_nC_k\, p^k (1-p)^{n-k}}, \quad (k = 0, 1, 2, \cdots, n)$$

となる。X の確率分布がこの式を満たすとき，確率 p に対する次数 n の**二項分布** (Binomial distribution) といい $B(n, p)$ で表す。例 2 と同様に考えると X の平均と分散は

$$\boxed{E[X] = np} \text{ (平均)}, \quad \boxed{V(X) = np(1-p)} \text{ (分散)}$$

である。4 ページの二項定理で $a = 1 - p$，$b = p$ とおくと $(a+b)^n = 1^n = 1$ より

$$\boxed{1 = {}_nC_0 p^0 (1-p)^n + {}_nC_1 p^1 (1-p)^{n-1} + {}_nC_2 p^2 (1-p)^{n-2} + \cdots + {}_nC_{n-1} p^{n-1} (1-p)^1 + {}_nC_n p^n (1-p)^0}$$

となるから二項分布の各確率の和が 1 であることがわかる。

問 コインを 7 回投げる。表の出る回数を X とすると，X の確率分布は二項分布 $B\left(7, \dfrac{1}{2}\right)$ である。このとき次の値を求めよ。

(1) $P(X = 0)$　　(2) $P(X = 3)$　　(3) $E[X]$　　(4) $V(X)$

< 確率変数の標準化 >

平均 m, 分散 v である確率変数 X に対し，$\sigma = \sqrt{v}$ を**標準偏差**という。
このとき

$$\boxed{X^* \;=\; \frac{X-m}{\sigma}}$$ (標準化された確率変数)

とおくと 19 ページ問の結果より

$$E[X^*] \;=\; \frac{1}{\sigma}(E[X]-m) \;=\; 0 \;\;,\;\; \mathrm{V}(X^*) \;=\; \left(\frac{1}{\sigma}\right)^2 \times v \;=\; 1$$

となる。つまり X^* は平均 0, 分散 1 の確率変数である。X^* を X の標準化された**確率変数**という。

例 1 サイコロを 20 回投げて 1 の出た回数を X とする。
X の分布は二項分布 $\mathrm{B}(20, \frac{1}{6})$ であるから

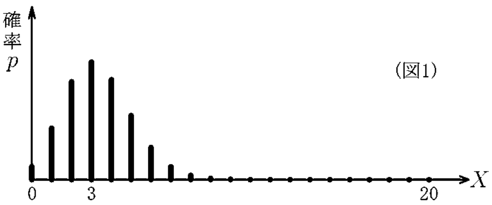
$$m = E[X] = 20 \times \frac{1}{6} = \frac{10}{3} \;,\;\; v = V(X) = 20 \times \frac{1}{6} \times \frac{5}{6} = \frac{25}{9}$$

より標準偏差は $\sigma = \sqrt{v} = \sqrt{\frac{25}{9}} = \frac{5}{3}$

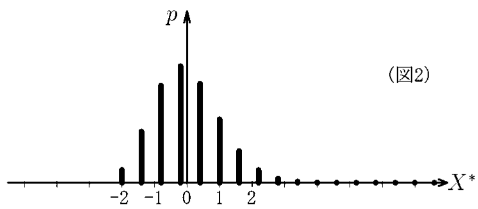
であるから，標準化された確率変数 X^* は

$$X^* = \frac{X-m}{\sigma} = \frac{X-\frac{10}{3}}{\frac{5}{3}} = \frac{3}{5}X-2$$

となる。図 1 は X の確率分布を棒グラフにしたもので，図 2 は X^* の確率分布である。



(図1)



(図2)

例 2 サイコロを 45 回投げて 1 の目の出た回数を X とする。
X の分布は二項分布 $\mathrm{B}(45, \frac{1}{6})$ だから

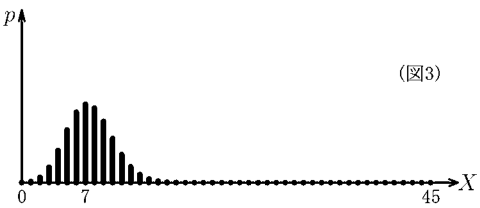
$$m = E[X] = 45 \times \frac{1}{6} = \frac{15}{2} \;,\;\; v = V(X) = 45 \times \frac{1}{6} \times \frac{5}{6} = \frac{25}{4}$$

より標準偏差は $\sigma = \sqrt{v} = \sqrt{\frac{25}{4}} = \frac{5}{2}$

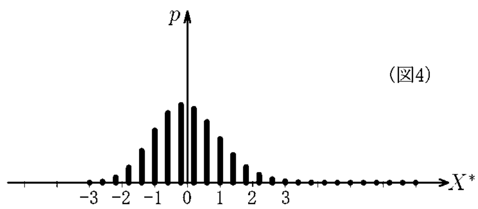
よって標準化された確率変数 X^* は

$$X^* = \frac{X-m}{\sigma} = \frac{X-\frac{15}{2}}{\frac{5}{2}} = \frac{2}{5}X-3$$

となる。図 3 は X の確率分布を棒グラフにしたもので，図 4 は X^* の確率分布である。



(図3)



(図4)

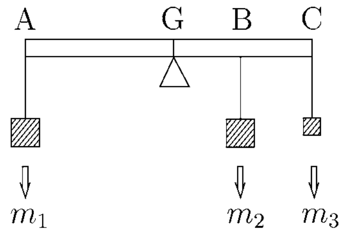
問 確率変数 X が以下の場合に，平均 m と標準偏差 σ を求め，標準化した確率変数 X^* を X で表せ。

(1) サイコロを 80 回投げて 1 の目の出た回数を X

(2) コインを 100 回投げて表の出た回数を X

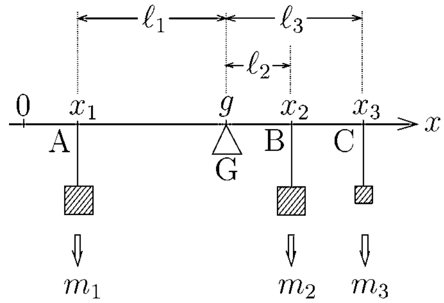
＜質量と重心１＞

例　細長い棒 AC に図１のように
おもり m_1, m_2, m_3 がかかっ
ているとする。棒自身のおもさ
を無視して重心 G の位置を
求めたい。



(図1)

この問題を図２のように
数直線上におもり m_1, m_2, m_3
がかかっていると考え,
各点の座標を x_1, x_2, x_3 として
重心の座標 g を求めたい。



(図2)

重心の意味から

(1)　$\ell_1 = g - x_1$,
　　$\ell_2 = x_2 - g$,　$\ell_3 = x_3 - g$

とおくと

(2)　$m_1 \times \ell_1 = m_2 \times \ell_2 + m_3 \times \ell_3$

が成り立つ。(2) 式に (1) を代入すると

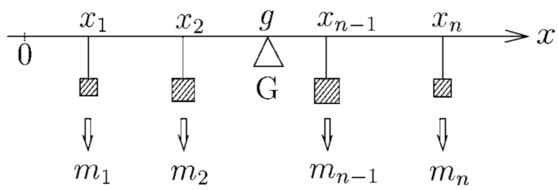
より
$$m_1 g - m_1 x_1 = (m_2 x_2 - m_2 g) + (m_3 x_3 - m_3 g)$$
$$(m_1 + m_2 + m_3) g = m_1 x_1 + m_2 x_2 + m_3 x_3$$

ここで全質量を $M = m_1 + m_2 + m_3$ とおくと

$$\boxed{g = \frac{1}{M}\{m_1 x_1 + m_2 x_2 + m_3 x_3\}}$$

が成り立つ。

問　数直線上に n 個のおもり
　m_1, m_2, $\cdots$, m_{n-1}, m_n
が図３のようにかかっているとき
重心 G の位置を全質量
　$M = m_1 + m_2 + \ \cdots + m_{n-1} + m_n$
と m_1, $\cdots$, m_n, x_1, $\cdots$, x_n を使って
表せ。

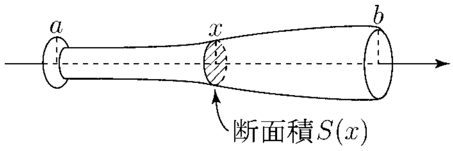


(図3)

＜ 質量と重心 2 ＞

例 野球のバットのような立体 (図 1) を考える。中心軸 (x 軸) に垂直な断面の断面積 $S(x)$ がわかっている場合，この立体の体積 V は

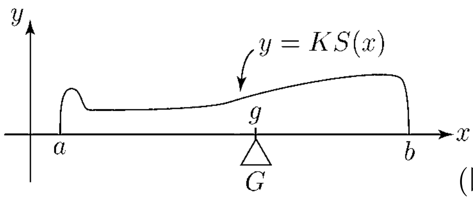
$$V=\int_a^b S(x)dx$$



(図 1)

であった。もしこのバットの材質が均一であれば，その質量 M は体積の定数倍 (K 倍) になると考えられるので

$$M=KV=\int_a^b KS(x)dx$$



(図 2)

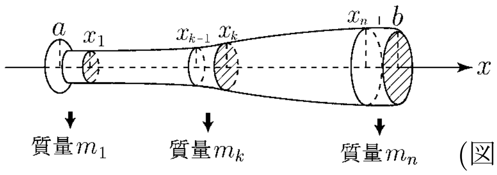
と表される。この場合被積分関数 $KS(x)$ をこの立体の**質量分布の密度関数**という。
この立体の重心 G の位置 g(図 2) を求めたい。
区間 $[a,b]$ を n 等分して，その分点を

$$a=x_0,\ x_1,\cdots,x_n=b$$

とおき，それぞれ

$$m_1,\ m_2,\ \cdots,m_n$$

の質量がかかっているとする (図 3)。



(図 3)

このとき $\Delta x=\dfrac{b-a}{n}$ とすると

$$m_k=K\int_{x_k-1}^{x_k} S(x)dx \fallingdotseq KS(x_k)\Delta x \qquad (k=1,2,\cdots,n)$$

である。前ページの結果より

$$g=\frac{1}{M}\left\{m_1x_1+m_2x_2+\cdots+m_nx_n\right\} \fallingdotseq \frac{1}{M}\left\{x_1KS(x_1)+\cdots+x_nKS(x_n)\right\}\Delta x$$

である。$n\to\infty$ とすれば定積分の区分求積法による定義から

$$\boxed{g=\frac{1}{M}\int_a^b xKS(x)dx} \quad (\textbf{重心})$$

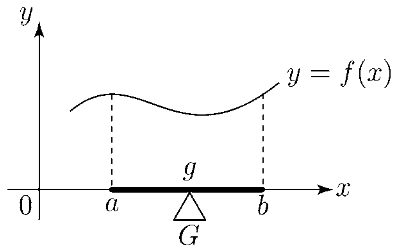
となる。

問 図 2 のように数直線の a から b までの区間に質量がある場合，質量分布の密度関係を $f(x)$ とすると，全質量 M は

$$M=\int_a^b f(x)dx$$

である。重心の座標 g を M と $f(x)$ を使って表せ。

< 質量と重心 3 >



数直線の区間 $[a, b]$ に質量があるとき，その質量分布の密度関数が $f(x)$ であれば，全質量 M と重心の座標 g は

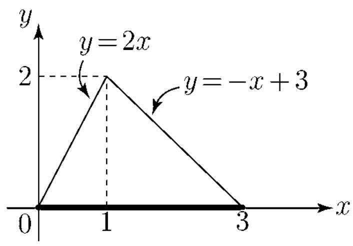
$$\boxed{M = \int_a^b f(x)dx \ \ , \ \ g = \frac{1}{M}\int_a^b xf(x)dx}$$

で表される。$f(x)$ を単に**密度関数**とか**重み関数**などという。

例 数直線上の区間 $[0, 3]$ に質量があり，その密度関数 $f(x)$ が

$$f(x) = \begin{cases} 2x & : \ 0 \leqq x \leqq 1 \\ -x+3 & : \ 1 \leqq x \leqq 3 \end{cases}$$

である場合，全質量 M と重心の座標 g は



$$M = \int_0^3 f(x)dx = \int_0^1 2x\,dx + \int_1^3 (-x+3)dx = \Big[x^2\Big]_0^1 + \Big[-\frac{x^2}{2} + 3x\Big]_1^3 = 3$$

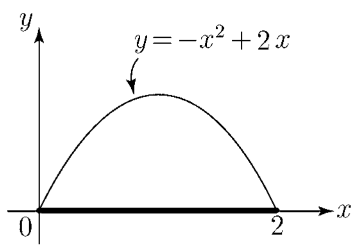
$$g = \frac{1}{M}\int_0^3 xf(x)dx = \frac{1}{3}\int_0^1 x \times 2x\,dx + \frac{1}{3}\int_1^3 x \times (-x+3)dx$$

$$= \frac{1}{3}\Big[\frac{2}{3}x^3\Big]_0^1 + \frac{1}{3}\Big[-\frac{x^3}{3} + \frac{3}{2}x^2\Big]_1^3 \ = \ \frac{4}{3}$$

問 数直線上の区間 $[0, 2]$ に質量があり，その密度関数 $f(x)$ が

$$f(x) = -x^2 + 2x \qquad (0 \leqq x \leqq 2)$$

である場合，全質量 M と重心の座標 g を求めよ。

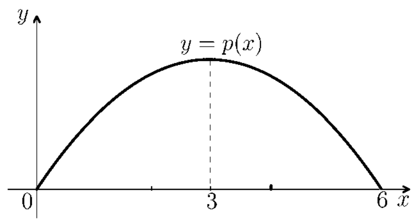


＜確率密度関数 1＞

非負な関数で，実数全体の積分が 1 になる関数を**確率密度関数**という。
すなわち

$$\boxed{p(x)\ :\ \text{確率密度関数}\ \Leftrightarrow\ p(x)\ \geqq\ 0\ ,\ \int_{-\infty}^{\infty} p(x)dx\ =\ 1}$$

例 1　定数 $k > 0$ に対し



$$p(x) = \begin{cases} kx(6-x) & : 0 \leqq x \leqq 6 \\ 0 & : \text{その他} \end{cases}$$

とおく。$p(x)$ が確率密度関数とすると

$$1 = \int_{-\infty}^{\infty} p(x)dx = \int_0^6 kx(6-x)dx = k\int_0^6 (6x - x^2)dx = k\left[3x^2 - \frac{x^3}{3}\right]_0^6 = 36k$$

より　　$k = \dfrac{1}{36}$

確率密度関数 $p(x)$ に対し，前ページの重心 g に相当する値を**平均**という。

$M = \displaystyle\int_{-\infty}^{\infty} p(x)dx = 1$ になるから，$p(x)$ の平均を m とすると

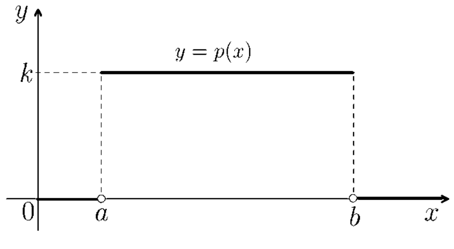
$$\boxed{m = \int_{-\infty}^{\infty} xp(x)dx}\ \text{(確率密度関数}\ p(x)\ \text{の\textbf{平均})}$$

例 2　例 1 の関数 $p(x)$ の平均を m とすると，

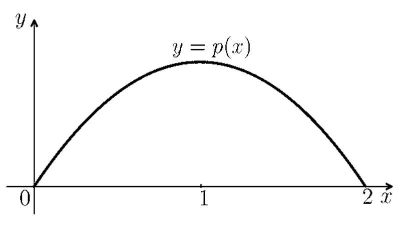
$$m = \int_{-\infty}^{\infty} xp(x)dx = \int_0^6 x\frac{1}{36}x(6-x)dx = \frac{1}{36}\int_0^6 (6x^2 - x^3)dx = \frac{1}{36}\left[2x^3 - \frac{x^4}{4}\right]_0^6 = 3$$

問　定数 k に対し，以下の関数 $p(x)$ が確率密度関数になるように k の値を求め，その平均 m を求めよ。

(1)　$p(x) = \begin{cases} k & : a \leqq x \leqq b \\ 0 & : \text{その他} \end{cases}$



(2)　$p(x) = \begin{cases} kx(2-x) & : 0 \leqq x \leqq 2 \\ 0 & : \text{その他} \end{cases}$



< 確率密度関数 2 >

平均が $m = \displaystyle\int_{-\infty}^{\infty} x\ p(x)\ dx$ である確率密度関数 $p(x)$ に対し

$$\boxed{v = \int_{-\infty}^{\infty} (x-m)^2 p(x) dx} \qquad \text{(分散)}$$

を $p(x)$ の **分散** といい，v で表す。

例　前ページ例 1 の $p(x)$ に対し，例 2 より $m = 3$ であるから

$$\begin{aligned} v &= \int_{-\infty}^{\infty} (x-m)^2 p(x) dx = \int_0^6 (x-3)^2\ \tfrac{1}{36} x(6-x) dx \\ &= \frac{1}{36} \int_0^6 (-x^4 + 12x^3 - 45x^2 + 54x) dx = \frac{1}{36} \left[-\frac{x^5}{5} + 3x^4 - 15x^3 + 27x^2 \right]_0^6 \\ &= \frac{1}{36} \left(-\frac{6^5}{5} + 3 \times 6^4 - 15 \times 6^3 + 27 \times 6^2 \right) = \frac{9}{5} \end{aligned}$$

問　前ページ問の確率密度関数 $p(x)$ に対し，分散 v を求めよ。

(1)　$p(x) = \begin{cases} \dfrac{1}{b-a} & : \quad a \leqq x \leqq b \\ 0 & : \quad \text{その他} \end{cases}$

(2)　$p(x) = \begin{cases} \dfrac{3}{4} x(2-x) & : \quad 0 \leqq x \leqq 2 \\ 0 & : \quad \text{その他} \end{cases}$

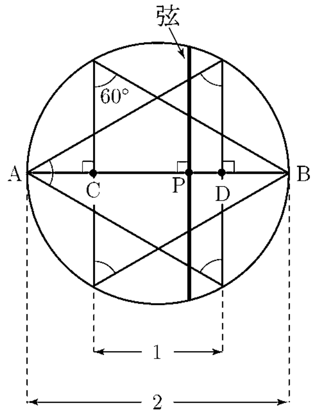
< Bertrand の問題 >

次の問題は J.Bertrand(1822-1900) により提出された。

> 問題「円に引いた任意の弦の長さが内接正三角形の一辺の長さ以上になる確率を求めよ。」

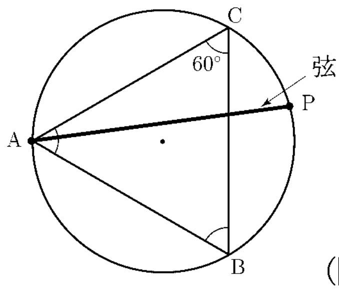
この問題に対し，以下の 2 つの解答例をあげる。

[解答 1]　図 1 の直径 AB に垂直な弦を考える。弦の長さが内接三角形の一辺より大きくなるのは図 1 の弦が線分 CD 上の点 P を通るときである。AB の長さを 2 とすれば CD の長さは 1 なので，求める確率は $\frac{1}{2}$ である。



(図1)

[解答 2]　図 2 の円周上の点 A を通る弦を考える。A を通る内接正三角形を ABC とする。A を通る弦の長さが AB = AC より長くなるのは弦のもう一端 P が弧 BC 上にあるときである。弧 BC の長さは円周の $\frac{1}{3}$ であるから，求める確率は $\frac{1}{3}$ である。



(図2)

さて，どちらが正しいか？· · · 実はこの問題には「正解」がない !!
「何が根元事象であるか」または「何が同様に確からしいか」という条件が問題の中にないので答は何通りでも作れる。

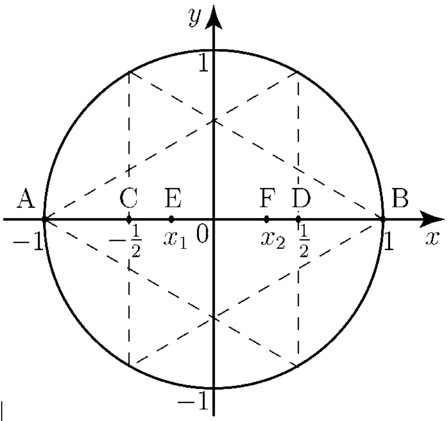
問　上の 2 つの解答例は “ 何 ” を「同程度に確からしい」と仮定しているか？

(1)　解答 1 の場合　　　　(2)　解答 2 の場合

＜ 連続分布 1 ＞

前ページの Bertrand の問題では弦は無限個引ける。確率の問題としては「事象」が「無限」にある。有限事象の場合は「各根元事象が同様に確からしい」と仮定するが,「無限事象」の場合は「何が同様に確からしい」かを明確に指定しないと確率は求まらない。

例　前ページ解答 1 では直径 AB 内の各点を通る (AB に垂直な) 弦を同様に確からしいと考えた。つまり AB 内の各点に弦が 1 つ 1 つ対応していて,「AB 内の各点が同様に確からしい」と仮定している。



確率の問題としてこの場合の全事象 U は「AB 内の各点を通る (AB に垂直な) 弦の全体」である。今右図のように座標平面上に 2 点 A$(-1,0)$, B$(1,0)$ をおく。この場合に全事象 U は { 線分 AB 内の各点の全体 } $=\{(x,0)\,:\,-1<x<1\}$ と同一視できる。

全事象 U を線分 AB と考えると,「AB 内の各点 ($= U$ の各事象) が同様に確からしい」とはどういうことであろうか?

イメージとしては「線分 AB に均一に質量がかかっている」と考えればよい。逆に「同様に確からしくない」場合のイメージとしては P25 のバットの例や, P26 の問のように「質量の分布が均一でない」場合を考えればよい。

「AB 内の各点が同様に確からしい」場合の確率は,「AB に均一な質量がかかっている」として,「確率」を「質量の割合」で定めればよい。

線分 AB を “ 同じ材質でできた直径一定の円柱状の棒 ” と考えると, “ 棒の重さ ” は “ 棒の長さ ” に比例するから,「確率」を「長さの割合」で考えればよい。

従って線分 CD 内を通る (AB に垂直な) 弦の確率は

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

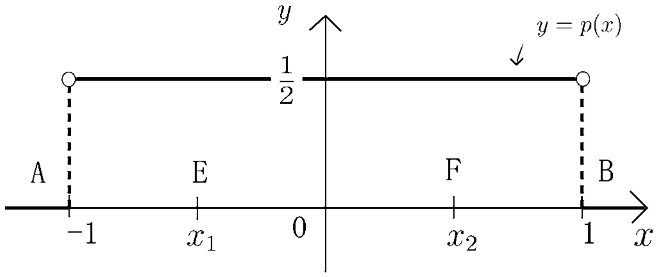
である。

問　上の例で線分 AB 内の任意の 2 点を E$(x_1,0)$, F$(x_2,0)$ $(-1<x_1<x_2<1)$ とする。線分 EF 内を通る (AB に垂直な) 弦の確率を求めよ。

＜連続分布 2＞

例 前ページの例で「線分 AB 内に均一な質量がかかっている」とした。
すなわち AB 間 $(-1 < x < 1)$ で質量分布の密度関数は定数になる。
AB の全質量を 1 とすると，質量分布の密度関数 $p(x)$ は $-1 < x < 1$ の範囲で定数である確率密度関数だから (P27問 (1) より)

$$p(x) = \begin{cases} \dfrac{1}{2} & : \quad -1 < x < 1 \\ 0 & : \quad \text{その他} \end{cases}$$



となる。

今全事象を $U = \{x : -1 < x < 1\}$ とし，X を弦の x 座標

$$X(x) = x \quad , \qquad (x \in U)$$

とおく。2 点 $\mathrm{E}(x_1, 0), \mathrm{F}(x_2, 0)(-1 < x_1 < x_2 < 1)$ の間に弦が通る確率を

$$P\Big(\{x \in U : x_1 < x < x_2\}\Big) = P(x_1 < X < x_2)$$

と略記すると，前ページ問より

$$P(x_1 < X < x_2) = \frac{x_2 - x_1}{2} = \int_{x_1}^{x_2} p(x)dx$$

と表せる。このような時「$p(x)$ を確率変数 X の確率密度関数である」とか,「確率変数 X は密度 $p(x)$ の分布に従う」などという。

問 1 上の例の場合に以下の確率を求めよ。

(1) $P(U) = P(-1 < X < 1) =$ 　　　　　　　　, (2) $P(-1 < X < x_1) =$

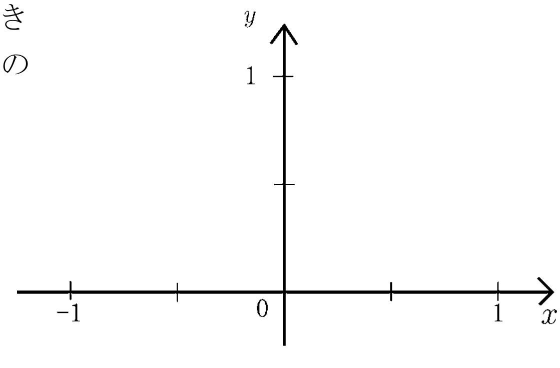
(3) $P(x_1 < X < 1) =$

(4) $P(X = x_1) = P(-1 < X < 1) - P(-1 < X < x_1) - P(x_1 < X < 1) =$

(5) $P(X = x_2) =$

(6) $P(x_1 \leqq X \leqq x_2) = P(X = x_1) + P(x_1 < X < x_2) + P(X = x_2) =$

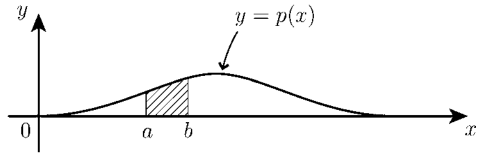
問 2 上の例の場合に $F(x) = P(X \leqq x)$ とおく。
$x \leqq -1$ のとき $F(x) = 0$ であり，$x \geqq 1$ のとき $F(x) = 1$ である。$-1 < x < 1$ のとき $F(x)$ を x の式で表し，右図に $y = F(x)$ のグラフを書け。



< 連続分布 3 >

確率変数 X と定数 $a, b (a < b)$ に対し,「 X が a より大きく b より小さい確率 」を $P(a < X < b)$ と略記する。今ある確率密度関数 $p(x)$ があって

$$(*) \quad \boxed{P(a < X < b) = \int_a^b p(x)dx} \quad \text{(連続分布の確率)}$$



で定まるとき,「 確率変数 X は **密度** $p(x)$ **の連続分布** に従う 」という。$y = p(x)$ のグラフが右図のような場合には $P(a < X < b)$ は図の斜線部分の面積である。

$p(x)$ は確率密度関数 $\left(p(x) \geqq 0 \,,\, \int_{-\infty}^{\infty} p(x)dx = 1 \right)$ であるから

$$(1) \quad \boxed{P(-\infty < X < \infty) = 1} \quad \text{(全事象の確率 1)}$$

であり，また任意の定数 a に対し,

$$(2) \quad \boxed{P(X = a) = \int_a^a p(x)dx = 0} \quad \text{(1 点 } a \text{ である確率 0)}$$

となる。(1) 式は確率変数 X がどんな場合でも成り立つ式であるが，(2) 式は連続分布特有の性質である。

(注 1)　サイコロ投げで出た目の数を X とすると $P(X = 1) = P(X = 2) = \cdots = P(X = 6) = \dfrac{1}{6}$ であり (2) 式は成り立たない。このような場合 X は**離散分布**という。

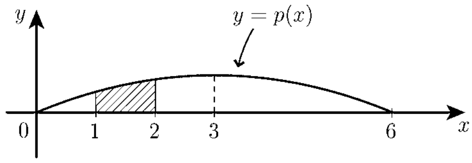
(注 2)　$F(x) = P(X \leqq x)$ を X の**分布関数**をいう。$(*)$ 式のような場合は，前ページ問 2 のように $y = F(x)$ のグラフが「連続」になるので「連続分布」という。注 1 のような場合は連続にならない。

確率変数 X が密度 $p(x)$ に従う場合 (=$(*)$ 式で確率が定まる場合) に，X の平均と分散は $p(x)$ の平均と分散で定める。

$$\boxed{m = \int_{-\infty}^{\infty} xp(x)dx} \quad \text{(平均)} \quad , \quad \boxed{v = \int_{-\infty}^{\infty} (x-m)^2 p(x)dx} \quad \text{(分散)}$$

例　X が密度

$$p(x) = \begin{cases} \frac{1}{36}x(6-x) & : \quad 0 \leqq x \leqq 6 \\ 0 & : \quad \text{その他} \end{cases}$$



に従う場合，P28 例より

平均 $m = 3$,　分散 $v = \frac{9}{5}$

である。また　$P(1 < X < 2) = \displaystyle\int_1^2 \frac{1}{36}x(6-x)dx = \frac{1}{36}\left[3x^2 - \frac{1}{3}x^3\right]_1^2 = \frac{5}{27}$

問　$p(x)$ が P28 問 (2) の場合に平均 m と分散 v を求め，$P(1 < X < 2)$ を計算せよ。

< 連続分布 4 >

確率変数 X が確率密度関数 $p(x)$ の分布に従う場合，すなわち

$$P(a < X < b) = \int_a^b p(x)dx \qquad \bigl(a\ ,\ b\ \ (a < b)\ \ \text{は任意の定数}\bigr)$$

である場合，その平均を

$$\boxed{E[X] \ = \ m \ = \int_{-\infty}^{\infty} xp(x)dx} \quad \textbf{(平均)}$$

と書くことにする。この平均値 m に対し，$(X-m)^2$ も確率変数であり，その平均は 19 ページと同様に

$$E[(X-m)^2] \ = \int_{-\infty}^{\infty} (x-m)^2 p(x)dx$$

となる。従って確率変数 X の分散は

$$\boxed{v \ = \ E[(X-m)^2] \ = \int_{-\infty}^{\infty} (x-m)^2 p(x)dx} \quad \textbf{(分散)}$$

と表される。ただし $m = E[X]$ である。
今定数 α , β に対し，

$$E[\alpha X + \beta] \ = \int_{-\infty}^{\infty} (\alpha x + \beta)p(x)dx \ = \ \alpha \int_{-\infty}^{\infty} xp(x)dx + \beta \int_{-\infty}^{\infty} p(x)dx$$

となるが $\displaystyle\int_{-\infty}^{\infty} xp(x)dx = E[X]$, $\displaystyle\int_{-\infty}^{\infty} p(x)dx = \ 1$ より

$$\boxed{E[\alpha X + \beta] \ = \ \alpha E[X] + \beta} \qquad (\alpha, \beta \text{ は定数})$$

が成り立つ。

例 平均 m，分散 v である確率変数 X に対し，

$$Y = 3X + 5$$

とおくと，Y も確率変数である。Y の平均を m_Y とすると

$$m_Y = E[Y] = E[3X+5] = 3E[X] + 5 = 3m + 5$$

となる。Y の分散を v_Y とすると

$$\begin{aligned} v_Y = E[(Y-m_Y)^2] &= E[\bigl(3X+5-(3m+5)\bigr)^2] = E[9(X-m)^2] \\ &= 9E[(X-m)^2] = 9v \end{aligned}$$

問 平均 m，分散 v である確率変数 X に対し

$$Y = \frac{X-m}{\sqrt{v}}$$

とおく。Y の平均 m_Y と分散 v_Y を求めよ。

< 連続分布 5 >

平均 m，分散 v である確率変数 X に対し，$\sigma = \sqrt{v}$ を X の**標準偏差**という。
このとき

$$\boxed{X^* = \frac{X-m}{\sigma}} \quad \text{(標準化された確率変数)}$$

とおくと，前ページ問の結果より X^* は平均 0，分散 1 の確率変数である。
X^* を X の**標準化された確率変数**という。

例 X が密度 $p(x)$ に従う確率変数で，平均 5，分散 9 であるとする。
このとき標準偏差は $\sigma = \sqrt{9} = 3$ である。標準化された確率変数 $X^* = \dfrac{X-5}{3}$ の密度関数を求めたい。今任意の定数 $a, b (a < b)$ に対して

$$P(a < X < b) = \int_a^b p(x)dx$$

である。よって

$$P(a < X^* < b) = P\left(a < \frac{X-5}{3} < b\right) = P(3a+5 < X < 3b+5) = \int_{3a+5}^{3b+5} p(x)dx$$

となる。ここで $y = \dfrac{x-5}{3}$ とおくと，$x = 3y+5$ であり，$\dfrac{dx}{dy} = 3$ より $dx = 3dy$ であるから置換積分法より

$$\int_{3a+5}^{3b+5} p(x)dx = \int_a^b p(3y+5)3dy$$

となる。よって

$$P(a < X^* < b) = \int_a^b 3p(3y+5)dy = \int_a^b 3p(3x+5)dx$$

であるから，X^* の密度関数は $3p(3x+5)$ である。

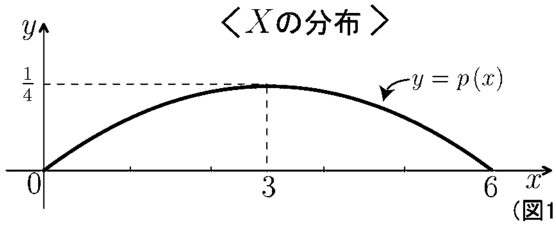
問 X が密度 $p(x)$ に従う確率変数で，平均 m，分散 v，標準偏差 $\sigma = \sqrt{v}$ であるとする。このとき標準化された確率変数 $X^* = \dfrac{X-m}{\sigma}$ の密度関数を求めよ。

＜連続分布 6 ＞

前のページの結果から

$$\boxed{X \text{ の密度が } p(x) \quad \Rightarrow \quad \tfrac{X-m}{\sigma} \text{ の密度は} \sigma p(\sigma x+m)}$$

である。



(図1)

例　確率変数 X の密度関数が

$$p(x)=\begin{cases}\frac{1}{36}x(6-x) & : 0 \leqq x \leqq 6 \\ 0 & : \text{その他}\end{cases}$$

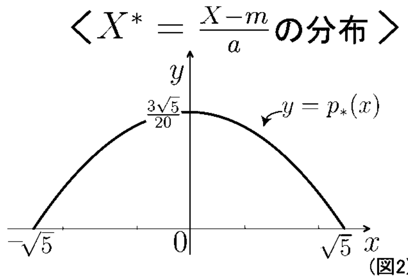
であるとき (図 1)，P28 例より

平均 $m=3$，分散 $v=\frac{9}{5}$，標準偏差 $\sigma=\sqrt{\frac{9}{5}}$

であった。従って標準化された確率変数

$$X^*=\frac{X-m}{\sigma}=\frac{X-3}{\sqrt{\frac{9}{5}}}$$

の密度関数は



(図2)

$$\begin{aligned}\sigma p(\sigma x+m) &= \sqrt{\tfrac{9}{5}} \times p\left(\sqrt{\tfrac{9}{5}}x+3\right) \\ &= \sqrt{\tfrac{9}{5}} \times \frac{1}{36}\left(\sqrt{\tfrac{9}{5}}x+3\right)\left(6-\left(\sqrt{\tfrac{9}{5}}x+3\right)\right) \\ &= \frac{3\sqrt{5}}{20}\left(1-\frac{x^2}{5}\right)\end{aligned}$$

で

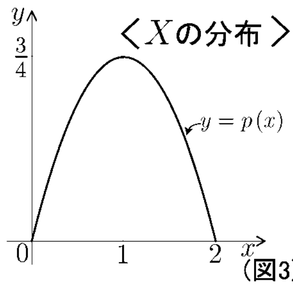
$X=0$ のとき $X^*=\frac{0-3}{\sqrt{\frac{9}{5}}}=-\sqrt{5}$

$X=6$ のとき $X^*=\frac{6-3}{\sqrt{\frac{9}{5}}}=\sqrt{5}$

より X^* の密度関数を $p_*(x)$ とすると

$$p_*(x)=\begin{cases}\frac{3\sqrt{5}}{20}\left(1-\frac{x^2}{5}\right) & : -\sqrt{5} \leqq x \leqq \sqrt{5} \\ 0 & : \text{その他}\end{cases}$$

となる (図 2)。



(図3)

問　確率変数 X の密度関数が

$$p(x)=\begin{cases}\frac{3}{4}x(2-x) & : 0 \leqq x \leqq 2 \\ 0 & : \text{その他}\end{cases}$$

とする (図 3)。P28 の結果を用いて，m と σ を求め，$X^*=\frac{X-m}{\sigma}$ の密度関数 $p_*(x)$ を求めよ。

＜ 正規分布 1 ＞

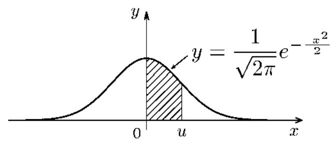
関数 $p(x) = \dfrac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$ は不定積分が求まらないが

$$\int_{-\infty}^{\infty} p(x)dx = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx = 1$$

であることはわかっている。この関数 $p(x) = \dfrac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$ を**標準正規分布の密度関数**という。

下の表はこの関数の定積分

$$\int_0^u \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx$$



の値を数値計算したもので,
正規分布表といわれる。

u	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.49379	.49396	.49413	.49430	.49446	49461	.49477	.49492	.49506	.49520
2.6	.49534	.49547	.49560	.49573	.49585	.49598	.49609	.49621	.49632	.49643
2.7	.49653	.49664	.49674	.49683	.49693	.49702	.49711	.49720	.49728	.49736
2.8	.49744	.49752	.49760	.49767	.49774	.49781	.49788	.49795	.49801	.49807
2.9	.49813	.49819	.49825	.49831	.49836	.49841	.49846	.49851	.49856	.49861
3.0	.49865	.49869	.49874	.49878	.49882	.49886	.49889	.49893	.49897	.49900
3.1	.49903	.49906	.49910	.49913	.49916	.49918	.49921	.49924	.49926	.49929
3.2	.49931	.49934	.49936	.49938	.49940	.49942	.49944	.49946	.49948	.49950
3.3	.49952	.49953	.49955	.49957	.49958	.49960	.49961	.49962	.49964	.49965
3.4	.49966	.49968	.49969	.49970	.49971	.49972	.49973	.49974	.49975	.49976
3.5	.49977	.49978	.49978	.49979	.49980	.49981	.49981	.49982	.49983	.49983
3.6	.49984	.49985	.49985	.49986	.49986	.49987	.49987	.49988	.49988	.49989
3.7	.49989	.49990	.49990	.49990	.49991	.49991	.49992	.49992	.49992	.49992
3.8	.49993	.49993	.49993	.49994	.49994	.49994	.49994	.49995	.49995	.49995
3.9	.49995	.49995	.49996	.49996	.49996	.49996	.49996	.49996	.49997	.49997

例 1

$$\int_0^{2.58} \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx = 0.49506$$

例 2

$$\int_{-1}^{1} \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx = 2\int_0^{1} \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx$$

$$= 2 \times 0.3413 = 0.6826$$

(注)　この関数は y 軸対称だから

$$\int_{-u}^{0} \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx = \int_0^{u} \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx$$

が成り立つ。

問　次の値を表から求めよ。

(1) $\displaystyle\int_{-1.96}^{1.96} \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx$

(2) $\displaystyle\int_{1}^{3} \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx$

(3) $\displaystyle\int_{-1}^{2} \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx$

< 正規分布 2 >

正の数 σ と実数 m に対し

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}$$

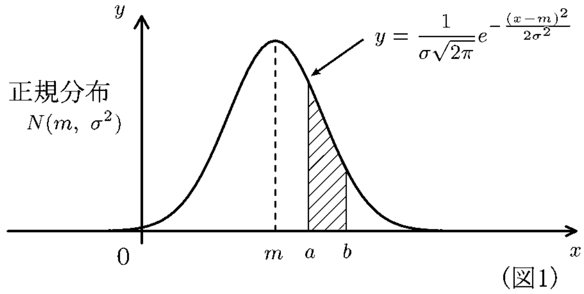
は平均 m，分散 σ^2 の確率密度関数である。確率変数 X がこの密度関数 $p(x)$ に従うとき，すなわち

$$P(a < X < b) = \int_a^b \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2} dx$$

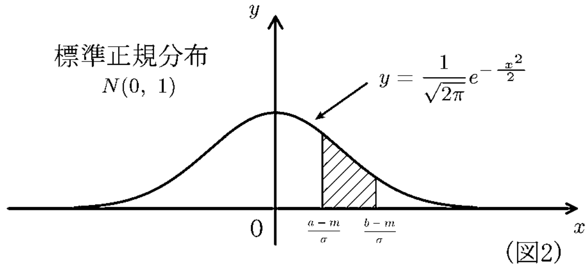
で表されるとき，X の分布を「平均 m，分散 σ^2 の**正規分布** (Normal Distribution)」といい，記号 $N(m, \sigma^2)$ で表す。X を標準化した確率変数を

$$X^* = \frac{X-m}{\sigma}$$

とおくと， X^* の分布は平均 0, 分散 1 の正規分布 $N(0,1)$ に従う。実際



(図1)



(図2)

$$P(a < X < b) = \int_a^b \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2} dx = \int_{\frac{a-m}{\sigma}}^{\frac{b-m}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du = P\left(\frac{a-m}{\sigma} < X^* < \frac{b-m}{\sigma}\right)$$

となる。$\left(\text{ここで } u = \dfrac{x-m}{\sigma} \text{ と変数変換している。}\right)$ 図 1 の斜線部分の面積が $P(a < X < b)$ であり，図 2 の斜線部分の面積と等しい。

例題 X が $N(30, 5^2)$ の正規分布に従うとき $P(25 \leqq X \leqq 40)$ を求めよ。

(解) X は連続分布なので $P(x = 25) = P(X = 40) = 0$ となる。$X^* = \dfrac{X-m}{\sigma} = \dfrac{X-30}{5}$ とおくと，X^* は標準正規分布に従うので

$$\begin{aligned} P(25 \leqq X \leqq 40) = P(25 < X < 40) &= P\left(\frac{25-30}{5} < \frac{X-30}{5} < \frac{40-30}{5}\right) \\ &= P(-1 < X^* < 2) = \int_{-1}^{2} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \end{aligned}$$

となる。前のページの表より

$$\int_{-1}^{2} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = \int_{-1}^{0} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx + \int_{0}^{2} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = 0.3413 + 0.4772 = 0.8185$$

となる。よって答は $P(25 \leqq X \leqq 40) = 0.8185$ である。

問 1 X が $N(10, 4^2)$ の正規分布に従うとき，次の確率を求めよ。

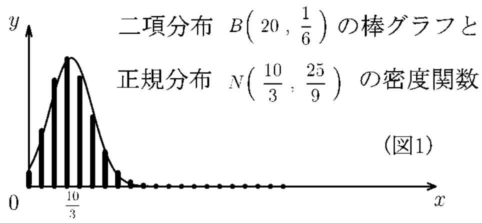
(1) $P(10 \leqq X \leqq 16)$, (2) $P(14 \leqq X \leqq 22)$, (3) $P(2 \leqq X \leqq 18)$

問 2 X が $N(m, \sigma^2)$ の分布に従うとき，次の確率を求めよ。

(1) $P(m - 1.96\sigma \leqq X \leqq m + 1.96\sigma)$ (2) $P(m - 2.58\sigma \leqq X \leqq m + 2.58\sigma)$

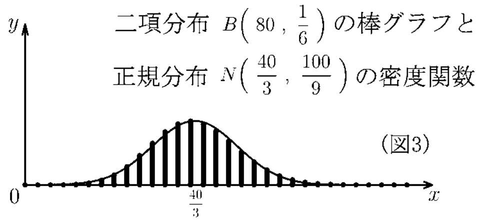
＜ 二項分布の正規近似 ＞

例 1 サイコロを 20 回投げて 1 の目の出た回数を X とする。X の分布は二項分布 $B\left(20\,,\ \frac{1}{6}\right)$ より 平均 $m = 20 \times \frac{1}{6} = \frac{10}{3}$，分散 $v = 20 \times \frac{1}{6} \times \frac{5}{6} = \frac{25}{9}$ である。図 1 は X の確率分布を棒グラフにしたものと，平均 $\frac{10}{3}$ 分散 $\frac{25}{9}$ の正規分布 $N\left(\frac{10}{3}\,,\ \frac{25}{9}\right)$ の密度関数を重ねて描いたもの。



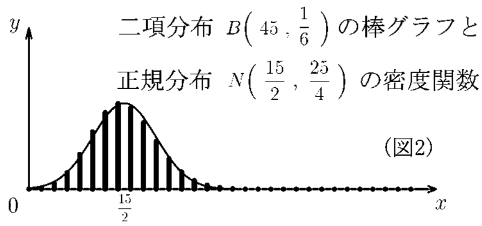
(図1)

例 2 サイコロを 45 回投げて 1 の目の出た回数を X とする。X の分布は二項分布 $B\left(45\,,\ \frac{1}{6}\right)$ より 平均 $m = 45 \times \frac{1}{6} = \frac{15}{2}$，分散 $v = 45 \times \frac{1}{6} \times \frac{5}{6} = \frac{25}{4}$ である。図 2 は X の確率分布を棒グラフにしたものと，平均 $\frac{15}{2}$ 分散 $\frac{25}{4}$ の正規分布 $N\left(\frac{15}{2}\,,\ \frac{25}{4}\right)$ の密度関数を重ねて描いたもの。



(図3)

例 3 サイコロを 80 回投げて 1 の目の出た回数を X とする。X の分布は二項分布 $B\left(80\,,\ \frac{1}{6}\right)$ より 平均 $m = 80 \times \frac{1}{6} = \frac{40}{3}$，分散 $v = 80 \times \frac{1}{6} \times \frac{5}{6} = \frac{100}{9}$ である。図 3 は X の確率分布を棒グラフにしたものと，平均 $\frac{40}{3}$ 分散 $\frac{100}{9}$ の正規分布 $N\left(\frac{40}{3}\,,\ \frac{100}{9}\right)$ の密度関数を重ねて描いたもの。



(図2)

例 4 サイコロを n 回投げて 1 の目の出た回数を X とする。X の分布は二項分布 $B\left(n\,,\ \frac{1}{6}\right)$ より

平均 $m = n \times \frac{1}{6} = \frac{n}{6}$，　分散 $v = n \times \frac{1}{6} \times \frac{5}{6} = \frac{5n}{36}$，　標準偏差 $\sigma = \sqrt{v} = \sqrt{\frac{5n}{36}}$ である。

n が大きくなると X の分布は平均 $\frac{n}{6}$ 分散 $\frac{5n}{36}$ の正規分布 $N\left(\frac{n}{6}\,,\ \frac{5n}{36}\right)$ に近づく。

すなわち正規分布で近似できる。$a < b$ に対し十分大きな n に対して以下のように近似できる。

$$P(a \leqq X \leqq b) \fallingdotseq \int_a^b \frac{1}{\sqrt{2\pi v}} e^{-\frac{(x-m)^2}{2v}}\,dx = \int_{\boldsymbol{a^*}}^{\boldsymbol{b^*}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\,dx$$

ただし

$$\boldsymbol{a^*} = \frac{a-m}{\sigma} = \frac{a-\frac{n}{6}}{\sqrt{\frac{5n}{36}}} \quad , \qquad \boldsymbol{b^*} = \frac{b-m}{\sigma} = \frac{b-\frac{n}{6}}{\sqrt{\frac{5n}{36}}}$$

一般に確率変数 X が二項分布 $\boldsymbol{B(\ n\ ,\ p\ )}$ に従うとき，n が十分大きければ，X の分布は平均 np，分散 $np(1-p)$ の正規分布 $\boldsymbol{N\left(\ np\ ,\ np(1-p)\ \right)}$ で近似できる。

$$\boxed{P(\ a \leqq X \leqq b\ ) \fallingdotseq \int_a^b \frac{1}{\sqrt{2\pi np(1-p)}} e^{-\frac{(x-np)^2}{2np(1-p)}}\,dx = \int_{\boldsymbol{a^*}}^{\boldsymbol{b^*}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\,dx} \quad \left(\boldsymbol{a^*} = \frac{a-np}{\sqrt{np(1-p)}} \quad , \quad \boldsymbol{b^*} = \frac{b-np}{\sqrt{np(1-p)}}\right)$$

(注)　より正確な近似を求めるときは，次の近似式を使う。

$$P(\ a \leqq X \leqq b\ ) \fallingdotseq \int_{\boldsymbol{(a-0.5)^*}}^{\boldsymbol{(b+0.5)^*}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\,dx \quad \left(\boldsymbol{(a-0.5)^*} = \frac{a-0.5-np}{\sqrt{np(1-p)}} \quad , \qquad \boldsymbol{(b+0.5)^*} = \frac{b+0.5-np}{\sqrt{np(1-p)}}\right)$$

問 　例 3 の場合に $P(10 \leqq X \leqq 20)$ の値を (注) の近似式を用いて近似せよ。

< 中心極限定理 >

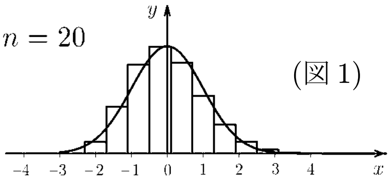
例 前ページ例 4 の場合 X の平均は $n\times\dfrac{1}{6}$ ，
標準偏差は $\sqrt{n\times\dfrac{5}{36}}$ より，標準化された確率変数は

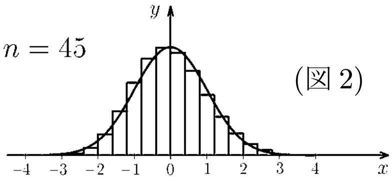
$$\boldsymbol{X^*}=\frac{X-(X\text{ の平均})}{(X\text{ の標準偏差})}=\frac{X-n\times\frac{1}{6}}{\sqrt{n\times\frac{5}{36}}}$$

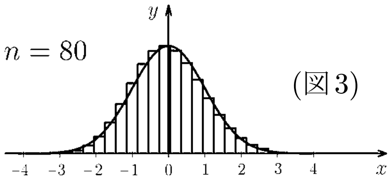
である。

図 1 は $n=20$ の場合の $\boldsymbol{X^*}$ の確率分布をヒストグラムの形にしたもので，各長方形の面積の和が 1 になるようにしたものである。このヒストグラムと標準正規分布 $N(0,1)$ の密度関数 $y=\dfrac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$ のグラフを重ねて書いてある。

図 2 は $n=45$ の場合であり，図 3 は $n=80$ の場合である。


$n=20$ (図 1)


$n=45$ (図 2)


$n=80$ (図3)

n が大きくなるにつれて $\boldsymbol{X^*}$ の分布は標準正規分布に近づく。すなわち

$$\lim_{n\to\infty}P(a\leqq \boldsymbol{X^*}\leqq b)=\int_a^b\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx$$

が任意の $a<b$ に対して成り立つ。

一方 22 ページ例 1 のように X_k を定めると $X_1,X_2,\cdots,X_n$ は独立で，しかも同じ分布 $\left(E[X_k]=\dfrac{1}{6}\,,\,V\big(X_k\big)=\dfrac{5}{36}\right)$ であり，

$$X=X_1+X_2+\cdots+X_n$$

と表される。従って上の極限式は次のようになる。$\left(\text{ただし}\boldsymbol{(X_1+X_2+\cdots+X_n)^*}=\dfrac{X_1+X_2+\cdots+X_n-n\times\frac{1}{6}}{\sqrt{n\times\frac{5}{36}}}\right)$

$$\lim_{n\to\infty}P\Big(a\leqq \boldsymbol{(X_1+X_2+\cdots+X_n)^*}\leqq b\Big)=\int_a^b\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx$$

一般に確率変数列 $X_1,X_2,\cdots,X_n,\cdots$ が独立で同じ確率分布をもつとする。
平均を $m=E[X_k]$，分散を $v=V\big(X_k\big)$ とすると，和 $X_1+X_2+\cdots+X_n$ を標準化した確率変数

(1) $$\boxed{\boldsymbol{(X_1+X_2+\cdots+X_n)^*}=\frac{X_1+X_2+\cdots+X_n-nm}{\sqrt{nv}}}$$ **(和の標準化)**

の分布は標準正規分布に近づく。すなわち

(2) $$\boxed{\lim_{n\to\infty}P\Big(a\leqq \boldsymbol{(X_1+X_2+\cdots+X_n)^*}\leqq b\Big)=\int_a^b\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx}$$ **(中心極限定理)**

が成り立つ (証明略)。これを**中心極限定理**という。(X_k の確率分布がどんな形でも成立する)
ここで (1) 式を (2) に代入し，$a=-M$ ，$b=M$ とおきかえると (2) 式は

(3) $$\lim_{n\to\infty}P\left(\left|\frac{X_1+X_2+\cdots+X_n}{n}-m\right|\leqq M\sqrt{\frac{v}{n}}\right)=\int_{-M}^{M}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx$$

と書きなおせる。ここで $n\to\infty$ のとき $M\sqrt{\dfrac{v}{n}}\to 0$ に注意してほしい。証明は略すが (3) 式から

(4) $$\lim_{n\to\infty}\frac{X_1+X_2+\cdots+X_n}{n}=m$$

が示される。(3) 式は「(4) 式の収束の速さが $\dfrac{1}{\sqrt{n}}$ に比例する」ことを意味している。

問 (1) 式と (2) 式から (3) 式を導け。

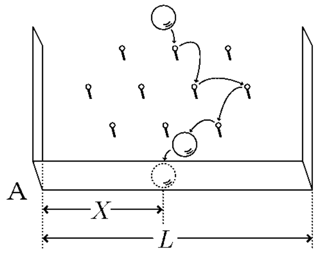
< 大数の法則 >

前ページで学んだように確率変数列 $X_1, X_2, \cdots, X_n, \cdots$ が独立で同じ確率分布をもつとき，その平均を $m = E[X_1] = E[X_2] = \cdots = E[X_n]$ とすれば,

$$\boxed{\lim_{n\to\infty} \frac{X_1 + X_2 + \cdots + X_n}{n} = m} \quad \text{(大数の法則)}$$

が成り立つ。これを**大数の法則**という。この法則を用いると，確率分布が未知である確率変数に対し，平均や分布を近似できる。

例 図 1 のようなパチンコ台の一部を使ってパチンコ玉を上からおとす。途中に水平な板 AB でパチンコ玉が止まるようにする。パチンコ玉の止まった位置 (A からの水平距離) を X とする。パチンコ玉は AB 間 (A と B の端は除く) の任意の位置に止まる可能性がある。つまり止まる位置 X は連続分布の確率変数と考えられる。しかし確率密度関数や平均値はわからない。

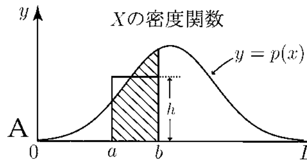


(図1)

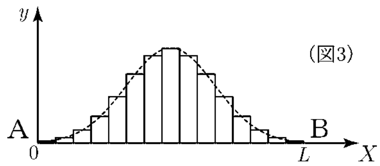
ところがその近似値は以下のようにして求められる。
このパチンコ玉を 1 回落とした玉の水平位置を X_1 とする。1 回目に落とした玉をとり除いて 2 回目の玉を落とす。
この試行を何回もくり返す。第 k 回目に落とした玉の位置を $X_k (k = 1, 2, \cdots, n, \cdots)$ とすると，$X_1, X_2, \cdots, X_n, \cdots$ は独立で同分布である。従って平均値 m は大数の法則より

$$m \fallingdotseq \frac{X_1 + X_2 + \cdots + X_n}{n}$$

で近似できる。



(図2)



(図3)

次に X の密度関数を近似的に求める。
仮に X の密度関数が $p(x)$ と仮定する。今任意の $a < b$ に対して積分値 $\int_a^b p(x)dx$ は図 2 の斜線部分の面積を表す。ここで

$$(\text{斜線部分の面積}) = \int_a^b p(x)dx = (b-a) \times h = (\text{長方形の面積})$$

となるように高さ h をえらぶ (図 2) と，h は

$$\boxed{\lim_{n\to\infty} \frac{(n\text{ 回のうち } a \text{ から } b \text{ の範囲に止まる回数})}{n(b-a)} = h} \quad \text{(分布密度に関する大数の法則)}$$

で求められる。高さ h は区間 $a \leqq x \leqq b$ における $p(x)$ の代表値と考えられる。
AB を等分し，各区間における $p(x)$ の代表値のヒストグラム (図 3) を作る。
等分を細かくすれば，密度関数 $p(x)$ が近似できる。

問 例の「分布密度に関する大数の法則」を以下の手順で証明せよ。

(1) Y_k を $a \leqq X_k \leqq b$ ならば $Y_k = 1$, そうでなければ $Y_k = 0$ とする。

$$P(Y_k = 1) = P(a \leqq X_k \leqq b) = \int_a^b p(x)dx$$

である。$E[Y_k]$ を求めよ。

(2) $Y_1 + Y_2 + \cdots + Y_n$ は何を表しているか説明せよ。

(3) $\displaystyle \lim_{n\to\infty} \frac{Y_1 + Y_2 + \cdots + Y_n}{n}$ を求めよ。

(4) $\displaystyle \lim_{n\to\infty} \frac{Y_1 + Y_2 + \cdots + Y_n}{n(b-a)}$ を求めよ。