目　　次

1

データの整理と表現

1-1 データの整理

連続変量と離散変量　身長や交通事故の件数などのように，ある集団において観測の対象となる特性を数量で表したものを**変量**という．変量には，身長や体重のように連続的な値をとる**連続変量**と，交通事故の件数や家族の人数のようにとびとびの値しかとらない**離散変量**とがある．

度数分布表　連続変量のデータを整理するときには，データをできるだけ同じ幅の区間に区切って分類する．このとき，各区間を**階級**といい，その中央の値を**階級値**という．また，各階級に分類されたデータの個数を**度数**という．階級の幅はデータ全体の傾向が読みとれるように，適当な大きさに選ぶ．階級または階級値に度数を対応させたものを**度数分布**といい，分類の結果でき上った表を**度数分布表**という．

離散変量のデータの分類は，その変量がとる値の個数を数えるだけでよいから簡単であるが，変量のとる値が多いときは，連続変量の場合のようにいくつかの階級を作って分類する．

累積度数分布表　各階級の度数の累計を**累積度数**といい，それらを表にまとめたものを**累積度数分布表**という．

相対度数分布表　各階級の度数をデータの総数で割った値を**相対度数**といい，これを表にしたものを**相対度数分布表**という．

度数分布表，相対度数分布表，累積度数分布表

階級	階級値	度数	相対度数	累積度数
$a_0 \sim a_1$	x_1	f_1	f_1/n	$F_1 = f_1$
$a_1 \sim a_2$	x_2	f_2	f_2/n	$F_2 = f_1 + f_2$
⋮	⋮	⋮	⋮	⋮
$a_{k-1} \sim a_k$	x_k	f_k	f_k/n	$F_k = f_1 + \cdots + f_k = n$
計		n	1	

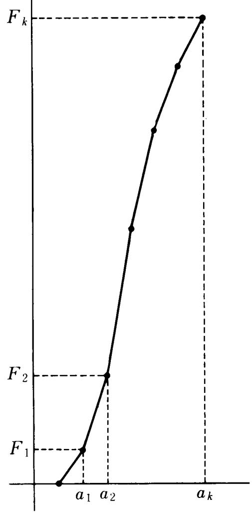
1-2　グラフによる表現

ヒストグラム　度数分布表を柱状グラフに表したもので，各階級の上に立てる長方形の面積は階級の度数に比例させる．

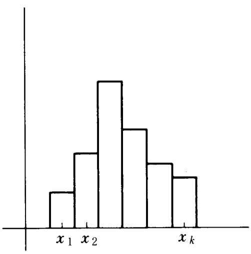
度数多角形　ヒストグラムの各長方形の上辺の中点を結んで得られる折れ線グラフ．

累積度数多角形　累積度数分布表をグラフに表したもので，階級の上方限界値と累積度数を座標にもつ点を結んで得られる．

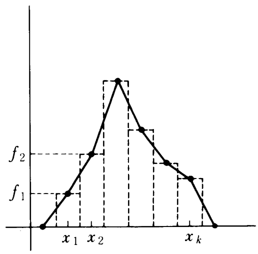
累積度数多角形



ヒストグラム



度数多角形



茎葉図（けい）　たとえば，10個のデータ（1.67, 1.82, 1.75, 1.63, 1.72, 1.79, 1.84, 1.60, 1.73, 1.75）が与えられたとき，各データを次のように“茎”に当たる（1.6, 1.7, 1.8）と“葉”に当たる3桁目の数に分解し，茎を縦に葉を横に記録することでデータを表現したものを**茎葉図**（または**幹葉図**）という．

茎葉図

茎	葉
1.6	0　3　7
1.7	2　3　5　5　9
1.8	2　4

1-3　代表値と散布度

（a）　生のデータからの場合.

n 個のデータを x_1, x_2, $\cdots$, x_n とする.

（b）度数分布表からの場合.

n 個のデータが k 個の階級に分類されているとして，階級値を x_1, x_2, $\cdots$, x_k, 度数を f_1, f_2, $\cdots$, f_k, 累積度数を F_1, F_2, $\cdots$, F_k とする.

1-3-1　代 表 値

データ全体の中心的傾向を表す値で，平均値，メジアン，モードなどがある.

平均（または**平均値**）

（a）　$\bar{x}=\dfrac{1}{n}\sum_{i=1}^{n}x_i$

（b）　$\bar{x}=\dfrac{1}{n}\sum_{i=1}^{k}x_if_i$

平均の簡便計算法　階級の幅を c, 仮平均を x_0 とする. ここで，変換 $u_i=\dfrac{x_i-x_0}{c}$ $(i=1, 2, \cdots, k)$ によって，x を u に換えれば

$$\bar{x}=x_0+c\bar{u}=x_0+c\frac{\sum_{i=1}^{k}u_if_i}{n}$$

メジアン

（a）　n 個のデータを大きさの順に並べたとき，n が奇数ならば小さい方から $\dfrac{n+1}{2}$ 番目の値を，n が偶数からば，中央にくる 2 つの値の平均を**メジアン**または**中央値**という.

（b）　この場合，メジアン M_e は次式より求められる.

$$M_e=a_{j-1}+c\frac{\frac{n}{2}-F_e}{f_j}$$

ここで，a_{j-1} はメジアンを含む階級の下方限界値，
f_j はメジアンを含む階級の度数，
F_e はメジアンを含む階級より前のすべての階級の度数の和，
すなわち $F_e=f_1+f_2+\cdots+f_{j-1}$,
c は階級の幅.

モード

（a）　データの中で最も多く現れている値を**モード**または**最頻値**という.

（b）　最大の度数をもつ階級の階級値.

1-3-2　散布度

データの散らばりを表す特性値で，標準偏差，分散，範囲などがある．

分散

（a）　$s^2=\dfrac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2$

$$=\frac{1}{n}\sum_{i=1}^{n}x_i^2-\bar{x}^2$$

（b）　$s^2=\dfrac{1}{n}\sum_{i=1}^{k}(x_i-\bar{x})^2f_i$

$$=\frac{1}{n}\sum_{i=1}^{k}x_i^2f_i-\bar{x}^2$$

分散の簡便計算法

$$s^2=c^2\left(\frac{1}{n}\sum_{i=1}^{k}u_i^2f_i-\bar{u}^2\right)$$

ただし，

$$u_i=\frac{x_i-x_0}{c},\qquad \bar{u}=\frac{1}{n}\sum u_if_i$$

標準偏差　　分散の正の平方根を**標準偏差**という．

（a）　$s=\sqrt{s^2}=\sqrt{\dfrac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2}$

$$=\sqrt{\frac{1}{n}\sum_{i=1}^{n}x_i^2-\bar{x}^2}$$

（b）　$s=\sqrt{s^2}=\sqrt{\dfrac{1}{n}\sum_{i=1}^{k}(x_i-\bar{x})^2f_i}$

$$=\sqrt{\frac{1}{n}\sum_{i=1}^{k}x_i^2f_i-\bar{x}^2}$$

標準偏差の簡便計算法

$$s=c\sqrt{\frac{1}{n}\sum u_i^2f_i-\bar{u}^2}$$

範囲　　データの中の最大値と最小値の差を**範囲**という．

四分位偏差　　n 個のデータを大きさの順に並べたとき，小さい方から $\left[\dfrac{n}{4}\right]+1$ 番目の値を**第 1 四分位数**，小さい方から $\left[\dfrac{3n}{4}\right]+1$ 番目の値を**第 3 四分位数**といい，それぞれ Q_1，Q_3 で表す．Q_3-Q_1 を**四分位範囲**，$\dfrac{1}{2}(Q_3-Q_1)$ を**四分位偏差**という．(注：$[x]$ は実数 x の整数部分)

例　題

例題 1　（離散データの度数分布）

次の数はある町での 30 日間の救急車の出動回数の記録である．このデータから度数分布表と累積度数分布表を作れ．また，(a) 平均値，(b) メジアン，(c) モードを求めよ．

2	6	1	1	2	0	3	0	2	3
3	0	4	2	1	1	0	3	4	0
5	1	4	5	1	0	5	0	2	0

解　データの分類の結果，

x	検数マーク	f	F
0	正下	8	8
1	正一	6	14
2	正	5	19
3	正	4	23
4	下	3	26
5	下	3	29
6	一	1	30
計		30	

したがって，度数分布表と累積度数分布表は

x	f	xf
0	8	0
1	6	6
2	5	10
3	4	12
4	3	12
5	3	15
6	1	6
計	30	61

x	F
0	8
1	14
2	19
3	23
4	26
5	29
6	30

(a)　平均値は，$\bar{x}=\dfrac{61}{30}\fallingdotseq \mathbf{2}$(回)．

(b)　累積度数分布表より小さい方から 15 番目と 16 番目の値はともに 2 で

あるから，メジアンは **2**(回).

（c）　最大度数は 8 だから，モードは **0**.

例題 2　（ヒストグラム・度数多角形・累積度数多角形）

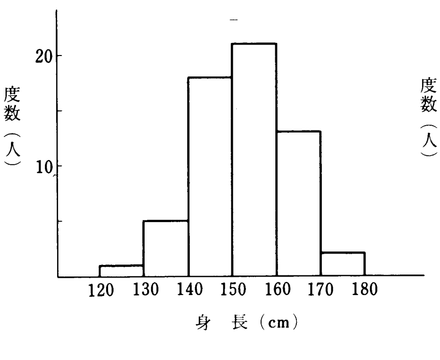
次の度数分布において，階級の真の限界，階級値および相対度数を与えよ．また，ヒストグラム，度数多角形，および累積度数多角形を図示せよ．

階級	度数
120-129	1
130-139	5
140-149	18
150-159	21
160-169	13
170-179	2
計	60

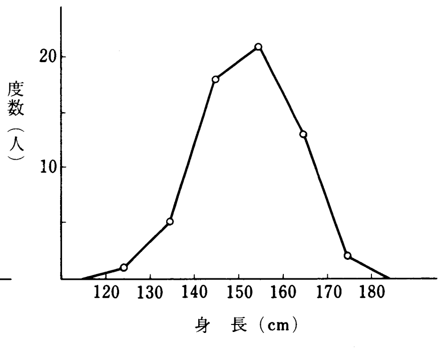
解

階級	階級の真の限界	階級値	度数	累積度数
120-129	119.5-129.5	124.5	1	1
130-139	129.5-139.5	134.5	5	6
140-149	139.5-149.5	144.5	18	24
150-159	149.5-159.5	154.5	21	45
160-169	159.5-169.5	164.5	13	58
170-179	169.5-179.5	174.5	2	60

ヒストグラム



度数多角形



まえがき

本書は大学1・2年次学生のための統計学の演習書である.

著者らの経験では，大学で初めて学ぶ統計学の講義内容は学生諸君にとって必ずしも理解しやすいとはいえないようである. 統計的考え方や統計学の基礎的概念を真に理解するためには，講議をきくだけでは十分ではなく，自らの力で問題を解いてみることが必要である.

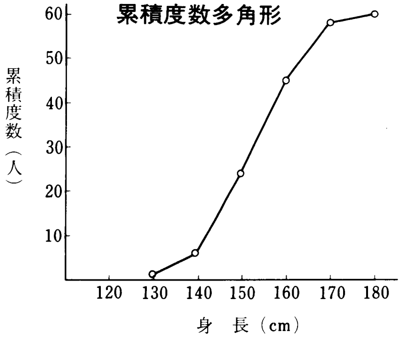
本書は各章の始めに公式や定義などを要項の形で簡潔にまとめ，その後に例題と演習問題が続く. 例題は理論的に難しい問題は避け，標準的でかつ興味ある実際的な問題を慎重に選び，すべてにていねいな解答を与えた. また，章末の演習問題も例題と同種なもので，これにも詳しい解答をつけている. 文系，理系を問わず広く学生諸君らが容易に理解できるよう，数学の予備知識としては高校の数学Ⅰ，基礎解析の程度とした.

この演習書は，講義の進行に合わせて予習，復習ができ，また試験などにも対応できるように配慮したつもりなので，必ずや読者諸氏の期待に沿うものと信ずる.

最後に，本書の出版に際して大変お世話になった培風館編集部の野原 剛，根岸誠一の両氏に心から御礼申しあげたい.

1988年12月

著　者



例題 3（茎葉図と度数分布）

男子学生 70 人の身長（cm）を測って，次の結果を得た.

169	176	169	169	167	170	172
163	182	173	175	163	168	173
171	164	168	162	167	163	177
159	161	167	164	176	164	159
172	169	164	174	165	177	175
157	175	154	169	155	176	183
167	168	159	168	171	162	173
172	170	173	171	166	168	168
171	164	161	165	168	164	156
174	175	168	161	180	181	166

このデータから，身長の

（a）茎葉図,

（b）度数分布,

（c）平均と標準偏差,

（d）メジアンとモード

を求めよ.

解（a）茎葉図

茎	葉
15	4 5 6 7 9 9 9
16	1 1 1 2 2 3 3 3 4 4 4 4 4 4 5 5 6 6 7 7 7 7 8 8 8 8 8 8 8 8 9 9 9 9 9
17	0 0 1 1 1 1 2 2 2 3 3 3 3 4 4 5 5 5 5 6 6 6 7 7
18	0 1 2 3

（b）　最大値は 183 で，最小値は 154 であるから，範囲は 183－154＝29．階級の幅を 5 cm にとれば，7 個の階級となる．

階級	検数マーク	階級の真の限界	度数分布 階級値	度数
150-154	一	149.5-154.5	152	1
155-159	正一	154.5-159.5	157	6
160-164	正正下	159.5-164.5	162	14
165-169	正正正正一	164.5-169.5	167	21
170-174	正正正	169.5-174.5	172	15
175-179	正下	174.5-179.5	177	9
180-184	下	179.5-184.5	182	4
			計	70

（c）　簡便計算法によって，平均と標準偏差を求める．階級の幅は $c=5$，仮平均を $x_0=167$ にとり

$$u=\frac{x-167}{5}$$

によって，x を u に変換する．

階級値 x	度数 f	u	uf	u^2f
152	1	-3	-3	9
157	6	-2	-12	24
162	14	-1	-14	14
167	21	0	0	0
172	15	1	15	15
177	9	2	18	36
182	4	3	12	36
計	70		16	134

$$\bar{u}=\frac{\sum u_i f_i}{n}$$
$$=\frac{16}{70}\fallingdotseq 0.229$$
$$\bar{x}=x_0+c\bar{u}$$
$$=167+5\times 0.229\fallingdotseq \mathbf{168.1}\ \text{(cm)}$$
$$s=c\sqrt{\frac{\sum u_i{}^2 f_i}{n}-\bar{u}^2}$$
$$=5\sqrt{\frac{134}{70}-\left(\frac{16}{70}\right)^2}\fallingdotseq \mathbf{6.8}\ \text{(cm)}$$

（d）　メジアンは公式より直接求める $n=70$，$c=5$，$a_{j-1}=164.5$，$F_e=1+6+14=21$，$f_j=21$ であるから

$$M_e=164.5+5\times\frac{35-21}{21}\fallingdotseq \mathbf{167.8}\ \text{(cm)}$$

モードは度数分布表より，**167**（cm）．

例題 4　**(階級の幅が等しくないときのヒストグラム)**

ある日，図書館で本を借りた人の年齢の分布は次のようであった．

年齢	8-12	13-20	21-60	61-64
人数	9	22	35	14

(a)　このデータをヒストグラムで表せ．

(b)　80人の年齢の平均と標準偏差を求めよ．

解　(a)　ここでの変量は年齢であるから階級の真の限界を定めるときに注意が必要である．たとえば，階級8-12は8歳ちょうどから13歳未満の人を含む(下表参照)この問題は階級の幅が等間隔でないから，ヒストグラムをかくとき，柱の面積が各度数に比例するようにしなければならない．これは，

$$\text{柱の高さ} \times \text{階級の幅} = \text{柱の面積} \propto \text{階級の度数}$$

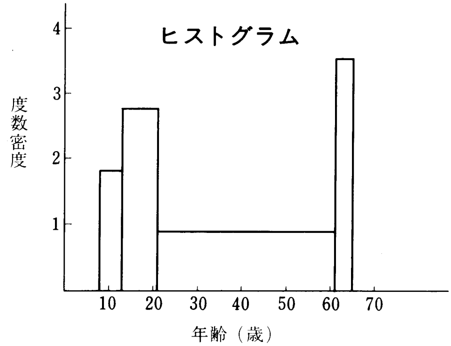
の関係より，

$$\text{柱の高さ} \propto \frac{\text{階級の度数}}{\text{階級の幅}}$$

よって，まず与えられた度数分布より $\dfrac{\text{階級の度数}}{\text{階級の幅}}$ (これを**度数密度**という)を求め，柱の高さをこの値に比例するようにとればよい．

階級	階級の真の限界	階級の幅	度数	度数密度
8-12	8-13	5	9	1.8
13-20	13-21	8	22	2.75
21-60	21-61	40	35	0.875
61-64	61-65	4	14	3.5

よって，



（b）　前の表より，

階級の真の限界	階級値 x	度数 f	xf	x^2f
8-13	10.5	9	94.5	992.25
13-21	17.0	22	374.0	6358.00
21-61	41.0	35	1435.0	58835.00
61-65	63.0	14	882.0	55566.00
	計	80	2785.5	121751.25

よって，　　平均　$\bar{x}=\dfrac{2785.5}{80}\fallingdotseq\mathbf{34.8}$（歳）

標準偏差　$s=\sqrt{\dfrac{121751.25}{80}-34.8^2}\fallingdotseq\mathbf{17.6}$（歳）

例題 5　（平均，標準偏差，メジアン，モード）

次の表は 1988 年全国高校野球選手権大会での全試合の勝敗の結果を示したものである．この表より

（a）　得点差の度数分布と累積度数分布を作れ．

（b）　得点差の平均と標準偏差を求めよ．

（c）　得点差のメジアンとモードを求めよ．

3-0	10-1	6-2	2-1
7-4	4-1	5-1	4-1
8-0	9-0	3-2	4-3
6-0	3-2	10-1	6-3
4-3	7-4	3-2	3-2
9-3	4-3	9-3	5-0
4-0	5-4	8-1	7-3
2-1	4-1	4-0	2-1
19-1	2-1	4-3	9-3
5-2	6-4	12-1	4-2
4-2	5-3	1-0	5-1
8-4	7-5	4-2	1-0

解　得点差 x を求めると

3	3	8	6	1	6	4	1	18	3	2	4
9	3	9	1	3	1	1	3	1	2	2	2
4	4	1	9	1	6	7	4	1	11	1	2
1	3	1	3	1	5	4	1	6	2	4	1

度数を f, 累積度数を F で表し, 次のようにデータを分類して, 度数分布と累積度数分布を作る.

（a）　度数分布　　累積度数分布

x	検数マーク	f	F
1	正正正	15	15
2	正一	6	21
3	正下	8	29
4	正丁	7	36
5	一	1	37
6	正	4	41
7	一	1	42
8	一	1	43
9	下	3	46
11	一	1	47
18	一	1	48
計		48	

x	f
1	15
2	6
3	8
4	7
5	1
6	4
7	1
8	1
9	3
11	1
18	1
計	48

x	F
1	15
2	21
3	29
4	36
5	37
6	41
7	42
8	43
9	46
11	47
18	48

（b）　度数分布表より

x	f	xf	x^2f
1	15	15	15
2	6	12	24
3	8	24	72
4	7	28	112
5	1	5	25
6	4	24	144
7	1	7	49
8	1	8	64
9	3	27	243
11	1	11	121
18	1	18	324
計	48	179	1193

よって,

$$\text{平均}\quad \bar{x}=\frac{179}{48}=\mathbf{3.7}\text{(点)}$$

$$\text{標準偏差}\quad s=\sqrt{\frac{1193}{48}-\left(\frac{179}{48}\right)^2}=\mathbf{3.3}\,(\text{点})$$

（c）　累積度数分布表より，小さい方から 24 番目と 25 番目の値はともに 3 だから，メジアンは **3**（点）.

モードは，度数分布表から **1**（点）.

例題 6　（メジアン）

次の数の平均値は 7 で，モードは 6 である．これら数のメジアンはいくつか.

$$8, 10, x, 6, y, z, 8, 21$$

解　モードが 6 であることから，x，y，z のうち少なくとも 2 つは 6 でなければならない．3 つとも 6 とすると，平均値が 7 にならないから 2 つが 6 である．いま，x，y を 6 とすると平均値は 7 だから

$$7=\frac{8+10+6+6+6+z+8+21}{8}=\frac{65+z}{8}\quad\Rightarrow\quad z=-9$$

よって，これらの数を大きさの順に並べると

$$-9, 6, 6, \underset{\uparrow}{6}, \underset{\uparrow}{8}, 8, 10, 21$$

ゆえに，メジアンは $\dfrac{6+8}{2}=\mathbf{7}$.

例題 7　（メジアンと四分位偏差）

高速道路のある地点で，走行中の車 400 台の時速（km/h）を測って，次の度数分布を得た.

階級値	30	35	40	45	50	55	60	65
度数	3	12	37	61	89	130	58	10

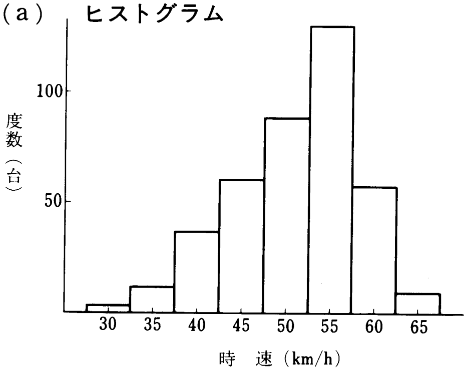
この結果を

（a）　ヒストグラム,

（b）　累積度数多角形

で表せ．求めた累積度数多角形からメジアンと四分位偏差を求めよ.

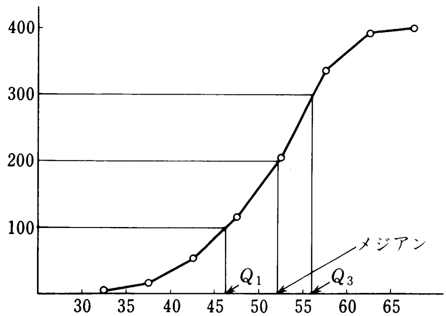
解　(a)　**ヒストグラム**



(b)　累積度数分布表は,

階級値	度数	階級の真の限界	累積度数
30	3	27.5-32.5	3
35	12	32.5-37.5	15
40	37	37.5-42.5	52
45	61	42.5-47.5	113
50	89	47.5-52.5	202
55	130	52.5-57.5	332
60	58	57.5-62.5	390
65	10	62.5-67.5	400
計	400		

累積度数多角形



この図からメジアンは **52**.

Q_1 は 46, Q_3 は 56. よって, 四分位偏差は $Q_3-Q_1=56-46=\mathbf{10}$.

例題 8　**（合併データの平均と標準偏差）**

公式

$$\sum_{i=1}^{n}(x_i-\bar{x})^2=\sum_{i=1}^{n}x_i^2-n\bar{x}^2$$

を証明せよ.

ある中学校の 3 年生は女子生徒 263 人と男子生徒 282 人からなり，女子の身長は平均 155.5 cm，標準偏差 4.0 cm で，男子の身長は平均 163.0 cm，標準偏差 4.4 cm である．この学校の 3 年生全員の身長の平均と標準偏差を求めよ.

解　公式の証明：

$$\begin{aligned}\sum(x_i-\bar{x})^2&=\sum(x_i^2-2x_i\bar{x}+\bar{x}^2)\\&=\sum x_i^2-2\bar{x}\sum x_i+n\bar{x}^2\\&=\sum x_i^2-2n\bar{x}^2+n\bar{x}^2\qquad\left(\textstyle\sum x_i=n\bar{x}\text{ より}\right)\\&=\sum x_i^2-n\bar{x}^2\end{aligned}$$

3 年生全員の身長の平均は

$$\bar{x}=\frac{263\times155.5+282\times163.0}{545}=\mathbf{159.4}\ (\mathrm{cm})$$

標準偏差の計算には上記の公式を使う．標準偏差を s とすると，この公式より

$$\sum x_i^2=n(\bar{x}^2+s^2)$$

女子生徒の場合，この値は

$$\sum x_i^2=263(155.5^2+4.0^2)=6363613.75$$

男子生徒の場合，この値は

$$\sum x_i^2=282(163.0^2+4.4^2)=7497917.52$$

よって，3 年生全員の身長の標準偏差は

$$\begin{aligned}s&=\sqrt{\frac{1}{n}\sum(x_i-\bar{x})^2}\\&=\sqrt{\frac{\sum x_i^2}{n}-\bar{x}^2}\\&=\sqrt{\frac{6363613.75+7497917.52}{545}-159.4^2}=\sqrt{25.64}=\mathbf{5.1}(\mathrm{cm})\end{aligned}$$

1 章の問題

1.1 次の各階級に対して，階級の真の限界，階級値，階級の幅を示せ．

(a) 20-29

(b) 150-, 180-

(c) 1.6-3.5

(d) (-8)-(-5)

1.2 (a) 数 $\{3, 5, 4, 1, 7\}$ の平均と標準偏差を求めよ．この結果を用いて，計算によらずに次の数の平均と標準偏差を求めよ．

(i) 13, 15, 14, 11, 17

(ii) $-2, 0, -1, -4, 2$

(iii) $a+3, a+5, a+4, a+1, a+7$

(iv) 0.3, 0.5, 0.4, 0.1, 0.7

(v) $3a+b, 5a+b, 4a+b, a+b, 7a+b$

(b) 次のデータから，平均，メジアン，モードを求めよ．ただし，$a<b<c<d<e$．

$$a, a, a, a, a, b, b, b, b, b, b, c, c, c, c, d, d, d, e, e$$

1.3 次の数はある電話交換局が 30 秒間隔で延べ 25 分間に受けた電話呼び出し数の記録である．

2	3	5	2	2	0	1	3	2	4
3	2	2	3	1	1	3	3	4	6
2	4	1	3	5	4	5	2	3	6
3	5	2	5	3	0	2	0	5	4
4	2	1	2	4	3	6	4	2	0

このデータから呼び出し数の度数分布を作り，度数多角形をかけ．また，呼び出し数の平均値，メジアン，モードを求めよ．

1.4 以下の数値は 100 人の生徒の IQ のデータである．55-64, 65-74, ･･･ を階級に選び，これらデータを分類して

(a) 度数分布表を作れ．

(b) IQ の平均と標準偏差を求めよ．

(c) IQ のメジアンを求めよ．

81	106	81	116	105	107	110	84	78	91
109	98	106	133	108	109	105	102	97	77
100	101	120	73	90	99	90	143	100	102
82	109	90	97	101	116	103	84	104	119
107	102	96	101	88	80	85	124	117	100
86	81	91	91	124	111	108	82	97	99
108	101	58	95	106	106	91	118	107	66
107	121	108	79	94	82	93	104	128	100
101	100	102	94	89	90	108	114	92	111
81	94	72	118	93	103	104	103	100	92

1.5　観測値 x, 5, y, 13 は大きさの順に並んでいる．これらの平均値は 7 で，メジアンは 6 である．分散を求めよ．

1.6　3 つのかごに，それぞれ 6 個，5 個，4 個のりんごが入っている．第 1 のかごのりんごの重さの平均値は 220 g で，第 2 のかごのりんごの重さの平均値は 280 g で，第 3 のかごのりんごの重さは

$$247\text{ g},\quad 250\text{ g},\quad 239\text{ g},\quad 264\text{ g}$$

である．第 3 のかごのりんごはどれもが第 1 のかごのりんごより重く，第 2 のかごのりんごより軽い．そのとき，これら 15 個のりんごの重さの平均値とメジアンとを求めよ．次に，これら 3 つのかごの重さを p, q, r とする．

$$p+q+r=240,\qquad p^2+q^2+r^2=20000$$

のとき，かごの重さの標準偏差を求めよ．

1.7　2 個のサイコロを同時に 100 回投げ，出た目の和 x を観測して次の度数分布を得た．x の平均と標準偏差を求めよ．

x	2	3	4	5	6	7	8	9	10	11	12	計
f	2	4	9	10	16	17	15	11	8	6	2	100

1.8　ある中学校の遅刻者 100 人の遅刻時間(分)の度数分布は次のようであった．

遅刻時間(分)	0-2	3-7	8-12	13-17	18-27	計
度数	40	37	13	5	5	100

(a) この分布のヒストグラムをかけ.

(b) 遅刻時間の平均と標準偏差を求めよ.

1.9 ある試験を受けた200人の生徒の得点の度数分布が, 次のように与えられたとき,

(a) この分布の平均と標準偏差を求めよ.

(b) 累積度数分布表を求め, 累積度数多角形をかけ.

階級	1-20	21-30	31-40	41-50	51-60	61-70	71-80	81-100
度数	4	15	43	55	39	31	10	3

1.10 4個の観測値の平均は3.13, 標準偏差は0.15である. これに, さらに6個の観測値

$$3.19 \qquad 2.86 \qquad 2.93 \qquad 3.15 \qquad 3.14 \qquad 3.21$$

が追加された. これら10個の観測値の平均と標準偏差を求めよ.

1.11 4個の数の平均は5, 分散は2で, 別な6個の数の平均は7, 分散は3である. これら10個の数の平均と分散を求めよ.

2

確　　率

2-1　標本空間と事象

同一条件の下での繰り返しが可能で，その結果が偶然に支配されるとみなせるような実験や観測を**試行**という．ある試行で起こるすべての結果の集合を**標本空間**といい，S で表す．S を構成する個々の結果を**標本点**，S の任意の部分集合を**事象**という．特に，1つの標本点のみからなる事象を**根元事象**という．

2-2　確率の定義

定義（1）　S を構成しているすべての標本点の起こりやすさが**同様に確からしい**とき，事象 E の起こる確率を

$$P(E)=\frac{n(E)}{n(S)}$$

で定義する．ここで，$n(E)$ と $n(S)$ はそれぞれ事象 E と標本空間 S に含まれる標本点の個数である．この $P(E)$ を事象 E の**確率**(または**数学的確率**)という．

定義（2）　n 回の試行で事象 E が r 回起こったとき，事象 E の起こる確率を，$n\to\infty$ のときの相対度数 $\frac{r}{n}$ の極限値で定義する．すなわち

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

この確率を事象 E の**経験的確率**(または**統計的確率**)という．n が大きいときの相対度数は経験的確率の近似値とみなされる．

確率の公理　次の3公理をみたす測度 $P(E)$ を事象 E の確率と定義する．

1°　$0\leqq P(E)\leqq 1$

2°　$P(S)=1$

3°　$E_i \cap E_j=\phi$ $(i \neq j)$ ならば，$P(E_1 \cup E_2 \cup \cdots)=P(E_1)+P(E_2)+\cdots$

2-3　事象と記号

和事象	$E_1 \cup E_2$	E_1 または E_2 が起こるという事象
積事象	$E_1 \cap E_2$	E_1 および E_2 が起こるという事象
余事象	$\overline{E}$	E が起こらないという事象
全事象	S	必ず起こる事象
空事象	ϕ	起こり得ない事象

2-4　確率の基本定理

(1)　任意の事象 E に対して

$$0 \leqq P(E) \leqq 1$$

特に，　$P(S)=1, \quad P(\phi)=0$

(2)　**余事象の確率**

$$P(\overline{E})=1-P(E)$$

(3)　**加法定理**

$$P(E_1 \cup E_2)=P(E_1)+P(E_2)-P(E_1 \cap E_2)$$

特に，E_1 と E_2 が互いに**排反**ならば

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

(4)　事象 E_1 が起こったという条件下で事象 E_2 の起こる**条件つき確率**を $P(E_2|E_1)$ で表し，次で定義する．

$$P(E_2|E_1)=\frac{P(E_1 \cap E_2)}{P(E_1)}$$

(5)　**事象の独立性**

$$E_1 \text{ と } E_2 \text{ が独立} \iff P(E_1 \cap E_2)=P(E_1)P(E_2)$$

(6)　**乗法定理**

$$\begin{aligned} P(E_1 \cap E_2)&=P(E_1)P(E_2|E_1)\\ &=P(E_2)P(E_1|E_2) \end{aligned}$$

特に，E_1 と E_2 が独立ならば

$$P(E_1 \cap E_2)=P(E_1)P(E_2)$$

2-5 ベイズの定理

n 個の事象 $E_1, E_2, \cdots, E_n$ は，$E_i \cap E_j = \phi$ $(i \neq j)$，$E_1 \cup E_2 \cup \cdots \cup E_n = S$ とする．B を任意の事象（ただし，$P(B) \neq 0$）とするとき，次が成り立つ．

ベイズの定理　$$P(E_k|B) = \frac{P(E_k)P(B|E_k)}{\sum_{j=1}^{n} P(E_j)P(B|E_j)} \quad (k=1, 2, \cdots, n)$$

例　題

例題 1　**（標本空間と確率）**

2 個のサイコロを転がすとき，次を求めよ．

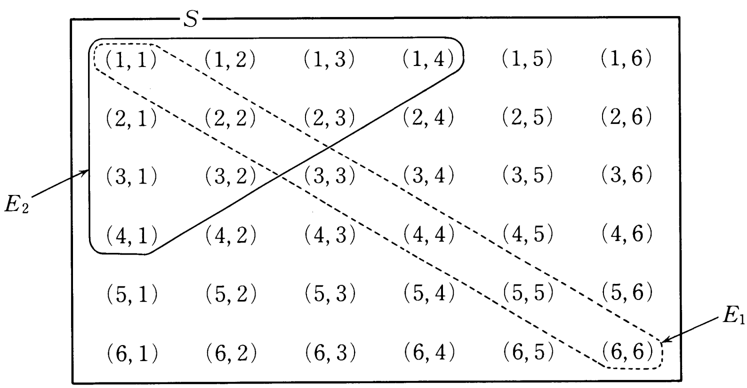
（a）　標本空間 S.

（b）　両方が同じ目を出す確率.

（c）　目の和が 5 以下である確率.

（d）　(b) または (c) が起こる確率.

解　（a）　標本空間 S は



同じ目が出る事象を E_1，目の和が 5 以下である事象を E_2 とすると，

$E_1 = \{(1, 1), (2, 2), (3, 3), (4, 4), (5, 5), (6, 6)\}$

$E_2 = \{(1, 1), (1, 2), (1, 3), (1, 4), (2, 1), (2, 2), (2, 3), (3, 1), (3, 2), (4, 1)\}$

上の図より，

$$n(S) = 36, \quad n(E_1) = 6, \quad n(E_2) = 10, \quad n(E_1 \cap E_2) = 2$$

各標本点は同様に確からしいから

（b）　$P(E_1)=\dfrac{n(E_1)}{n(S)}=\dfrac{6}{36}=\mathbf{\dfrac{1}{6}}$

（c）　$P(E_2)=\dfrac{n(E_2)}{n(S)}=\dfrac{10}{36}=\mathbf{\dfrac{5}{18}}$

（d）　$P(E_1\cup E_2)=P(E_1)+P(E_2)-P(E_1\cap E_2)=\dfrac{1}{6}+\dfrac{5}{18}-\dfrac{2}{36}=\mathbf{\dfrac{7}{18}}$

例題 2　**（経験的確率）**

次の表は新生児1万人が特定の年齢まで生き残ると期待される人数を示したものである．

年齢	0	10	20	30	40	50	60	70	80	90
生存者数	10000	9590	9510	9295	9067	8651	7590	5278	3625	1530

この表を用いて，次の確率を求めよ．

（a）　いま生まれた新生児がこの後10年以内に死ぬ．

（b）　いま生まれた新生児が60歳まで生きる．

（c）　いま生まれた新生児が30歳から40歳までの間に死ぬ．

（d）　いま40歳の人が今後10年以内に死ぬ．

解　（a）　新生児1万人のうち，10歳までに死ぬのは$10000-9590=410$（人）だから，求める確率は$\dfrac{410}{10000}=\mathbf{0.041}$.

（b）　新生児1万人のうち，60歳まで生き残る人は7590人いるので，求める確率は$\dfrac{7590}{10000}=\mathbf{0.759}$.

（c）　30歳から40歳までに死亡する人は$9295-9067=228$（人）いるから，求める確率は$\dfrac{228}{10000}=\mathbf{0.0228}$.

（d）　これは条件つき確率で，40歳まで生きた人9067人のうち，50歳までに死亡する人は$9067-8651=416$（人）だから，求める確率は$\dfrac{416}{9067}=\mathbf{0.0459}$.

例題 3 （ベン図と確率）

ある市では A, B, C の 3 紙の新聞が販売されている．この市で世帯を対象に新聞購読調査を行った結果，次のことがわかった．

20%が A を購読．

16%が B を購読．

14%が C を購読．

8%が A, B の 2 紙を購読．

5%が A, C の 2 紙を購読．

4%が B, C の 2 紙を購読．

2%が A, B, C の 3 紙を購読．

このとき，3 紙のうち

（a） 少なくとも 1 紙，

（b） 1 紙のみ，

（c） A 紙のみ

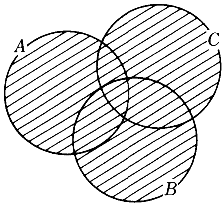
を購読する世帯の割合を求めよ．

解 ベン図を使って解く．題意から

$$P(A)=0.20,\ P(B)=0.16,\ P(C)=0.14,\ P(A\cap B)=0.08,$$
$$P(A\cap C)=0.05,\ P(B\cap C)=0.04,\ P(A\cap B\cap C)=0.02$$

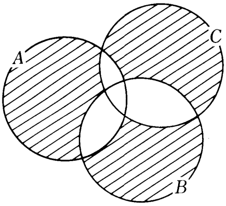
（a）

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



$$=0.20+0.16+0.14-0.08-0.05-0.04+0.02$$
$$=\mathbf{0.35}$$

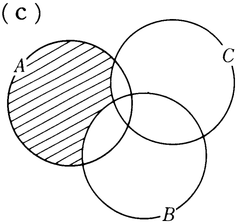
（b）



左図より

$$P(A\cup B\cup C)-P(A\cap B)-P(A\cap C)$$
$$-P(B\cap C)+2P(A\cap B\cap C)$$
$$=0.35-0.08-0.05-0.04+2\times 0.02$$
$$=\mathbf{0.22}$$

（c）



左図より

$P(A)-P(A\cap B)-P(A\cap C)+P(A\cap B\cap C)$

$=0.20-0.08-0.05+0.02$

$=\mathbf{0.09}$

例題 4　（ベン図と確率）

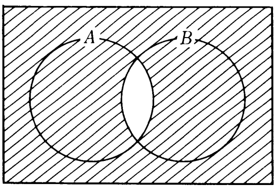
$P(A)=a$, $P(B)=b$, $P(A\cap B)=c$ とするとき，次の事象の確率を a, b, c で表せ．

（a）　$\overline{A}\cup\overline{B}$　　（b）　$\overline{A}\cap\overline{B}$　　（c）　$\overline{A}\cup B$

（d）　$\overline{A}\cap B$　　（e）　$\overline{A\cup B}$　　（f）　$\overline{A\cap B}$

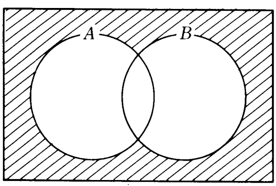
解　このような問題を解くときには，必要に応じてベン図を使うと便利である．

（a）



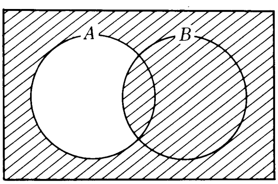
$$P(\overline{A}\cup\overline{B})=P(\overline{A\cap B})$$
$$=1-P(A\cap B)$$
$$=\mathbf{1-c}$$

（b）



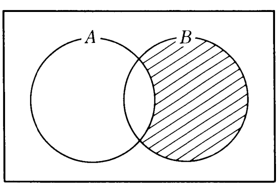
$$P(\overline{A}\cap\overline{B})=P(\overline{A\cup B})$$
$$=1-P(A\cup B)$$
$$=1-\{P(A)+P(B)-P(A\cap B)\}$$
$$=\mathbf{1-a-b+c}$$

（c）

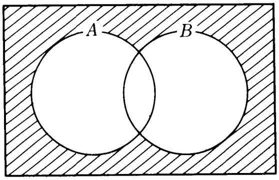


$$P(\overline{A}\cup B)=\mathbf{1-a+c}$$

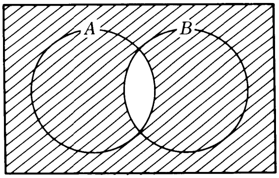
（d）



$$P(\overline{A}\cap B)=\mathbf{b-c}$$

(e) 

$$P(\overline{A \cup B}) = P(\overline{A} \cap \overline{B})$$
$$= \mathbf{1 - a - b + c} \qquad ((\mathrm{b})\text{と同値})$$

(f) 

$$P(\overline{A \cap B}) = P(\overline{A} \cup \overline{B})$$
$$= \mathbf{1 - c} \qquad ((\mathrm{a})\text{と同値})$$

例題 5　**(確率の計算)**

ある選挙では夫婦のうち夫が投票する確率は 0.5, 妻が投票する確率は 0.6, 夫が投票したことがわかったとき, その妻が投票する確率は 0.9 であるという. この選挙で

(a)　夫妻がともに投票する確率,

(b)　夫妻のうち, 少なくとも一方が投票する確率,

(c)　妻が投票したことがわかったとき, その夫が投票する確率

を求めよ.

解　夫が投票する確率を $P(A)$, 妻が投票する確率を $P(B)$ とする.

(a)　題意より $P(A)=0.5$, $P(B|A)=0.9$ であるから

$$P(A \cap B) = P(A)P(B|A)$$
$$= 0.5 \times 0.9 = \mathbf{0.45}$$

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

$$= 0.5 + 0.6 - 0.45$$
$$= \mathbf{0.65}$$

(c)　$P(A|B) = \dfrac{P(A \cap B)}{P(B)} = \dfrac{0.45}{0.6} = \mathbf{0.75}$

例題 6　**(確率の計算)**

A と B は互いに排反な事象で, $P(A)=0.2$, $P(B)=0.8$ である. いま, $P(C|A)=0.4$, $P(C|B)=0.5$ のとき, $P(A|C)$ を求めよ.

解
$$0.4=P(C|A)=\frac{P(C\cap A)}{P(A)}=\frac{P(C\cap A)}{0.2}\Rightarrow P(C\cap A)=0.08$$
$$0.5=P(C|B)=\frac{P(C\cap B)}{P(B)}=\frac{P(C\cap B)}{0.8}\Rightarrow P(C\cap B)=0.40$$
A と B が排反ならば，$C\cap A$ と $C\cap B$ も排反であるから
$$\begin{aligned}P(C)&=P((C\cap A)\cup(C\cap B))\\&=P(C\cap A)+P(C\cap B)\\&=0.08+0.40=0.48\end{aligned}$$
よって，
$$P(A|C)=\frac{P(C\cap A)}{P(C)}=\frac{0.08}{0.48}=\frac{1}{6}\fallingdotseq\mathbf{0.167}$$

例題 7　（確率の計算）

2 次方程式
$$x^2+ax+b=0$$
の 2 つの係数をサイコロの 2 回の投げで決めるとき，得られた 2 次方程式が

（a）　実数解，（b）　有理解

をもつ確率を求めよ．

解　サイコロの 2 回の投げで決まる (a, b) のすべての組合せは，次の 36 通りである．

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

（a）　2 次方程式で実数解をもつのは，判別式が $D=a^2-4b\geqq 0$ を満たすときで，これを満たす (a, b) の組合せは

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

の 19 組だから，求める確率は $\mathbf{\frac{19}{36}}$．

（b）　2次方程式が有理解をもつのは，判別式 $D=a^2-4b$ が0か，有理数の平方であればよい．

これを満たす (a, b) の組合せは，

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

の7組だから，求める確率は $\boldsymbol{\dfrac{7}{36}}$.

例題 8　**（確率の計算）**

52枚のトランプ札をよく切って4人のプレーヤーにそれぞれ13枚ずつ配るとき

（a）　特定のプレーヤーの手にある種類の札が一枚も含まれない確率,

（b）　特定のプレーヤーの手にエースが4枚とも配られる確率

を求めよ.

（c）　ゲームを繰り返して，特定のプレーヤーがすべてのエースを少なくとも1回受けとる確率を0.5以上とするためには何回のゲームが必要か.

解　（a）　特定のプレーヤーの手にある種類の札，たとえばスペードがこない確率は，このプレーヤーの手にスペードを除いた39枚の中から13枚がくればよいから

$$\frac{{}_{39}C_{13}}{{}_{53}C_{13}}=\frac{\dfrac{39!}{13!26!}}{\dfrac{52!}{13!39!}}\fallingdotseq 0.01279$$

ハート，クラブ，ダイヤについても同じことが考えられるから，求める確率は

$$0.01279\times 4\fallingdotseq \mathbf{0.051}$$

（b）　エースが4枚配られる確率は，13枚中4枚がエースで，あとの9枚がエース以外の48枚からとられればよいから

$$\frac{{}_4C_4\times{}_{48}C_9}{{}_{53}C_{13}}=\frac{\dfrac{48!}{9!39!}}{\dfrac{52!}{13!39!}}=\mathbf{0.0026}$$

（c）　1回のゲームで，特定のプレーヤーに4枚のエースが配られる確率を p，配られない確率を q とする．題意より

$$\sum_{r=1}^{n}{}_nC_r p^r q^{n-r}\geqq 0.5$$

を満たす n を求めればよい．この不等式の左辺は $1-q^n$ に等しく，(b) より，$q=1-p=1-0.0026=0.9974$ であるから

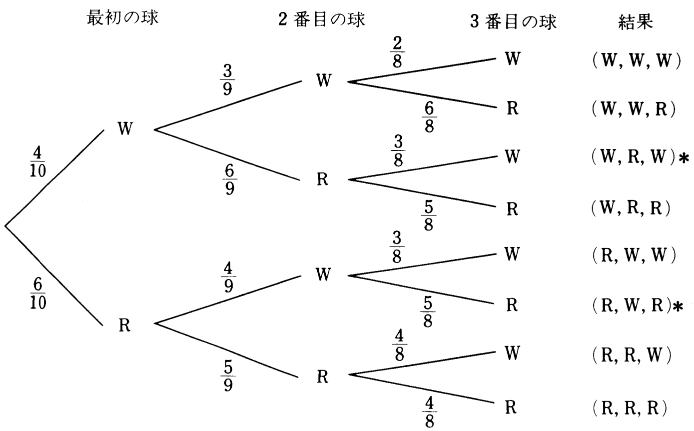
$$(0.9974)^n \leqq 0.5$$
$$n \log 0.9974 \leqq \log 0.5$$
$$n \geqq 266.24$$

よって，**267 回以上．**

例題 9　**（樹形図と確率）**

4 個の白球と 6 個の赤球を含む箱から，非復元抽出で 1 個ずつ無作為に 3 個をとるとき，同じ色の球が続けて出ない確率を求めよ．

解　白球を W，赤球を R で表し，次の樹形図を作る．



同じ色の球が続けて出ない場合は＊印のついた (W, R, W) と (R, W, R) の場合で，これらは互いに排反であるから，求める確率は

$$P\{(\mathrm{W, R, W}) \cup (\mathrm{R, W, R})\}$$
$$=P\{(\mathrm{W, R, W})\}+P\{(\mathrm{R, W, R})\}$$
$$=\frac{4}{10}\times\frac{6}{9}\times\frac{3}{8}+\frac{6}{10}\times\frac{4}{9}\times\frac{5}{8}=\mathbf{\frac{4}{15}}$$

例題 10 （ベイズの定理）

3つの袋をA, B, Cとする．各袋は10個の球を含み，Aは白球3個と赤球7個，Bは白球5個と赤球5個，Cは白球7個と赤球3個を含む．いま，1つの袋をランダムに選び，選ばれた袋から1球をとり出したら白球が出た．この球がA, B, Cのそれぞれから出たという確率を求めよ．

解　白球がとり出されるという事象を W とすると，与えられた情報より

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

$$P(W|A)=\frac{3}{10},\quad P(W|B)=\frac{5}{10},\quad P(W|C)=\frac{7}{10}$$

よって，ベイズの定理から

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

$$=\frac{\frac{1}{3}\times\frac{3}{10}}{\frac{1}{3}\times\frac{3}{10}+\frac{1}{3}\times\frac{5}{10}+\frac{1}{3}\times\frac{7}{10}}=\mathbf{\frac{1}{5}}$$

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

$$=\frac{\frac{1}{3}\times\frac{5}{10}}{\frac{1}{3}\times\frac{3}{10}+\frac{1}{3}\times\frac{5}{10}+\frac{1}{3}\times\frac{7}{10}}=\mathbf{\frac{1}{3}}$$

$$P(C|W)=1-P(A|W)-P(B|W)$$

$$=1-\frac{1}{5}-\frac{1}{3}=\mathbf{\frac{7}{15}}$$

2章の問題

2.1　(a)　$P(A\cup B)=\frac{7}{8}$, $P(A\cap B)=\frac{1}{4}$, $P(\overline{A})=\frac{5}{8}$ のとき，

(i) $P(A)$, (ii) $P(\overline{B})$, (iii) $P(A\cap\overline{B})$

を求めよ．

(b)　$P(A)=\frac{1}{2}$, $P(B)=\frac{1}{3}$, $P(A\cap B)=\frac{1}{5}$ のとき，

(i) $P(A|B)$, (ii) $P(B|A)$, (iii) $P(A\cup B)$, (iv) $P(\overline{A}|B)$

を求めよ．

2.2 2 つの袋 A と B がある．A は 2 個の赤球と 8 個の白球を含み，B は 3 個の赤球と 7 個の白球を含む．サイコロを投げ，1 または 2 の目が出たら A から 1 球をとり出し，1，2 以外の目が出たら，B から 1 球をとり出す．サイコロの投げで，1 または 2 の目が出る事象を X，赤球がとり出される事象を Y とするとき，次の確率を求めよ．

(a) $P(X)$, (b) $P(Y)$, (c) $P(X \cap Y)$, (d) $P(Y|X)$,
(e) $P(\overline{X} \cap Y)$

2.3 第 1 の箱は a 個の赤球と b 個の青球と c 個の白球を含み，第 2 の箱は，p 個の赤球と q 個の青球と r 個の白球を含む．いま，硬貨を投げ，おもてが出たら第 1 の箱から，うらが出たら第 2 の箱から，非復元抽出で 2 個の球を無作為にとり出す．どちらの箱も 40 個の球を含むとき，とり出した 2 個が同じ色である確率は

$$\frac{1}{3120}(a^2+b^2+c^2+p^2+q^2+r^2)-\frac{1}{39}$$

となることを示せ．

2.4 あるゲームを A，B，C の 3 人で競う．A が B に勝つ確率は 0.60 で，B が C に勝つ確率は 0.40 で，C が A に勝つ確率は 0.55 である．“クジ”で不戦勝を決めて勝負を競うとき，このゲームで A が優勝する確率はいくらか．

2.5 3 人の競技者 A，B，C が 5 個の白球と 5 個の黒球の入った“ツボ”から 1 度に 1 個ずつ非復元抽出で球がなくなるまで順次球をとり出す．最初に白球をとり出したものを勝ちとし，ゲームは A → B → C の順で行うとき，A，B，C がこのゲームにそれぞれ勝つ確率を求めよ．

2.6 (a) 52 枚のトランプ札から非復元抽出で 3 枚をランダムにとるとき，次の確率を求めよ．

(i) 3 枚とも同じ種類である．

(ii) 3 枚とも同じ数である．

(iii) 3 枚が同じ種類か同じ数のいずれかである．

(b) 3 個のサイコロを投げる．そのとき，出た目の和が

(i) 8 または 9 になる，

(ii) 完全平方になる

確率を求めよ．

2.7 赤球3個と白球4個を含む袋から2個の球を無作為にとり出し，得られた球の色をみた後，それらを袋にもどす．次に，再度2個の球をとり出し，球の色をみる．そのとき，次の確率を求めよ．

(a) 最初のとり出しで赤球が2個出て，次のとり出しで白球が2個出る．

(b) とり出された4個の球が赤球2個と白球2個よりなる．

(c) とり出された球が4個とも同じ色である．

2.8 ある円の内部でランダムに1点を選ぶとき，その点が円周までの距離より円の中心までの距離の方に近い確率は $\dfrac{1}{4}$ であることを示せ．

(a) この円内で何個かの点を逐次選ぶとき，4番目の点がはじめて円周より円の中心に近い方に入る確率を求めよ．

(b) 円周より円の中心に近い点をはじめて得る確率を0.90以上にするには，少なくとも何個の点を選ばねばならないか．

2.9 ある養鶏場にはA，B 2品種のめんどりが飼われている．ここで産み落される卵の80%は品種Aによるものである．品種Aのめんどりが産む卵の大きさはサイズ1が30%，サイズ2が45%，サイズ3が25%である．品種Bのめんどりの場合，これらの比率はそれぞれ35%，40%，25%である．卵の色(褐色と白色)は両品種とも大きさとは無関係であり，品種Aの40%と品種Bの30%は褐色である．このとき，次の確率を求めよ．

(a) 品種Bのめんどりの産んだ卵がサイズ1で褐色である．

(b) 産み落された卵がサイズ1で白色である．

(c) 白色の卵がサイズ1である．

2.10 ある病気に対する成人の罹患率は1%であることが知られている．この病気の発見に有効と見られる検査法が開発された．この病気にかかっている成人患者の90%はこの検査法に陽性反応を示し，病気にかかっていない成人は0.5%が同じ陽性反応を示した．無作為に選ばれたある成人がこの検査法で陽性反応を示したとき，この人が本当にその病気にかかっている確率を求めよ．

3

確 率 分 布

3-1 確 率 変 数

試行の結果に対応していろいろな値をとる変数を**確率変数**という．確率変数には整数などの値をとる**離散型確率変数**と，ある区間内の任意の実数値をとり得る**連続型確率変数**とがある．以下においては，確率変数を表すのに X，Y 等の大文字を用いる．

3-2 確 率 分 布

離散型確率分布　離散型確率変数 X のとる値を $x_1, x_2, \cdots, x_k$，X がこれらの値をとる確率を $p_1, p_2, \cdots, p_k$ とする．すなわち，

$$P(X=x_i)=p_i \qquad (i=1, 2, \cdots, k)$$

$$\sum_{i=1}^{k} p_i=1$$

このとき，$x_1, x_2, \cdots, x_k$ と $p_1, p_2, \cdots, p_k$ との対応関係を X の**確率分布**（または**離散型確率分布**）という．確率分布は通常，次のような表で示すことが多い．

表 1　X の確率分布

x	x_1	x_2	$\cdots$	x_k	計
$P(X=x)$	p_1	p_2	$\cdots$	p_k	1

連続型確率分布　連続型確率変数 X が区間 $[a, b]$ 内の値をとる確率が

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

で表されるとき，$f(x)$ を X の**確率密度関数**または単に**密度関数**という．

$f(x)$ の性質　(i)　すべての x に対して　$f(x) \geqq 0$

(ii)　$\int_{-\infty}^{\infty} f(x)\,dx = 1$

分布関数　$F(x) = P(X \leqq x) = \int_{-\infty}^{x} f(t)\,dt$

で定義される関数 $F(x)$ を確率変数 X の**累積分布関数**または単に**分布関数**という.

$F(x)$ の性質　(i)　$F(-\infty) = 0$,　$F(+\infty) = 1$

(ii)　$F(x)$ は非減少関数

(iii)　$\dfrac{dF(x)}{dx} = f(x)$

3-3 結合確率分布

結合確率分布　2 つの離散型確率変数 X, Y がそれぞれ x_i, y_j をとる確率を

$$P(X = x_i,\ Y = y_j) = p_{ij} \qquad (i = 1, 2, \cdots, k\,;\, j = 1, 2, \cdots, m)$$

とするとき, これを 2 次元確率変数 (X, Y) の**結合確率分布**（または**同時確率分布**）という.

結合確率分布も 1 次元の確率分布の場合と同じく, 次のような表で示すことが多い.

表 2　(X, Y) の結合確率分布

x ＼ y	y_1	y_2	$\cdots$	y_m	$P(X=x)$
x_1	p_{11}	p_{12}	$\cdots$	p_{1m}	$p_{1\cdot}$
x_2	p_{21}	p_{22}	$\cdots$	p_{2m}	$p_{2\cdot}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
x_k	p_{k1}	p_{k2}	$\cdots$	p_{km}	$p_{k\cdot}$
$P(Y=y)$	$p_{\cdot 1}$	$p_{\cdot 2}$	$\cdots$	$p_{\cdot m}$	1

周辺分布

$$P(X = x_i) = p_{i\cdot} = \sum_{j=1}^{m} p_{ij} \qquad (i = 1, 2, \cdots, k)$$

$$P(Y = y_j) = p_{\cdot j} = \sum_{i=1}^{k} p_{ij} \qquad (j = 1, 2, \cdots, m)$$

このとき, x_i と $p_{i\cdot}$ の対応を X の**周辺分布**, 同様に y_j と $p_{,j}$ の対応を Y の**周辺分布**という.

確率変数の独立性　2 つの離散型確率変数 X と Y が独立であるとは

すべての i, j に対して,

$$P(X=x_i,\ Y=y_j)=P(X=x_i)P(Y=y_j)$$

すなわち,

$$p_{ij}=p_{i\cdot}\,p_{\cdot j}$$

が成り立つときをいう.

3-4　確率変数の期待値と分散

期待値　確率変数 X の期待値を $E(X)$ で表し, 次で定義する.

$$E(X)=\begin{cases}\displaystyle\sum_{i=1}^{k} x_i p_i & (X\text{ が離散型のとき})\\ \displaystyle\int_{-\infty}^{\infty} xf(x)\,dx & (X\text{ が連続型のとき})\end{cases}$$

期待値は平均値ともいい, そのときは $E(X)$ の代りに μ で表す. $\mu\equiv E(X)$.

確率変数の関数の期待値　X の関数 $g(X)$ の期待値を次で定義する.

$$E[g(X)]=\begin{cases}\displaystyle\sum_{i=1}^{k} g(x_i)p_i & (X\text{ が離散型のとき})\\ \displaystyle\int_{-\infty}^{\infty} g(x)f(x)\,dx & (X\text{ が連続型のとき})\end{cases}$$

期待値に関する性質

(i)　$E(aX+b)=aE(X)+b$　　(a, b は定数)

(ii)　$E(X+Y)=E(X)+E(Y)$

　　$E(X-Y)=E(X)-E(Y)$

(iii)　X と Y が独立ならば

$$E(XY)=E(X)E(Y)$$

分散　確率変数 X の分散を $V(X)$ で表し, 次で定義する.

$$\begin{aligned}V(X)&=E[\{X-E(X)\}^2]\\ &=E[(X-\mu)^2]=\begin{cases}\displaystyle\sum_{i=1}^{k} (x_i-\mu)^2 p_i & (X\text{ が離散型のとき})\\ \displaystyle\int_{-\infty}^{\infty} (x-\mu)^2 f(x)\,dx & (X\text{ が連続型のとき})\end{cases}\end{aligned}$$

分散はしばしば σ^2 で表される. $\sigma^2\equiv V(X)$.

分散に関する性質

(i)　$V(X)=E(X^2)-[E(X)]^2$

(ii)　$V(aX+b)=a^2V(X)$　　($a,\ b$ は定数)

(iii)　X と Y が独立ならば

$$V(X+Y)=V(X)+V(Y)$$
$$V(X-Y)=V(X)+V(Y)$$

標準偏差　分散の正の平方根

$$\sigma=\sqrt{V(X)}=\sqrt{E[(X-\mu)^2]}$$

を**標準偏差**という.

標準化　X の平均値を μ, 標準偏差を σ とするとき

$$Z=\frac{X-\mu}{\sigma}$$

を X の**標準化変量**という. Z の平均は 0 で, 分散は 1 である.

$$E(Z)=0, \quad V(Z)=1$$

例　題

例題 1　(**期待値**)

昭和 55 年発行のジャンボ宝くじの賞金金額は 1 ユニット (1000 万本) 当たり, 次のようであった.

1 等		30,000,000 円	10 本
	組違い	150,000 円	90 本
2 等		10,000,000 円	10 本
	組違い	80,000 円	90 本
3 等		5,000,000 円	10 本
	組違い	50,000 円	90 本
4 等		1,000,000 円	300 本
5 等		100,000 円	100 本
6 等		10,000 円	1,000 本
7 等		300 円	2,000,000 本

宝くじ 1 本の獲得賞金額 X は離散型確率変数である. X の期待値を求めよ.

解　はずれくじの本数は

$10{,}000{,}000-(10+90+10+90+10+90+300+100+1{,}000+2{,}000{,}000)$

$=7{,}998{,}300$

よって, X の期待値の定義より

$$E(X)=30{,}000{,}000\times\frac{10}{10{,}000{,}000}+150{,}000\times\frac{90}{10{,}000{,}000}+\cdots$$

$$+300\times\frac{2{,}000{,}000}{10{,}000{,}000}+0\times\frac{7{,}998{,}300}{10{,}000{,}000}=\mathbf{139.52}\ (\text{円})$$

例題 2 （離散型確率分布の平均と分散）

X の確率分布が

x	-1	1	2	4
$P(X=x)$	0.1	0.4	0.3	0.2

のとき，

（a） $E(X)$, $E(X^2)$, $V(X)$ を求めよ．

（b） $Y=2X-3$ のとき，$E(Y)$, $V(Y)$ を求めよ．

解 （a） $E(X)=(-1)\times 0.1+1\times 0.4+2\times 0.3+4\times 0.2=\mathbf{1.7}$

$E(X^2)=(-1)^2\times 0.1+1^2\times 0.4+2^2\times 0.3+4^2\times 0.2=\mathbf{4.9}$

$V(X)=E(X^2)-\{E(X)\}^2=4.9-1.7^2=\mathbf{2.01}$

（b） $Y=2X-3$ の確率分布は

y	-5	-1	1	5
$P(Y=y)$	0.1	0.4	0.3	0.2

であるから

$E(Y)=(-5)\times 0.1+(-1)\times 0.4+1\times 0.3+5\times 0.2=\mathbf{0.4}$

$V(Y)=E(Y^2)-\{E(Y)\}^2$

$=(-5)^2\times 0.1+(-1)^2\times 0.4+1^2\times 0.3+5^2\times 0.2-0.4^2=\mathbf{8.04}$

例題 3 （確率変数の平均と分散）

X が平均 μ, 分散 σ^2 をもつとき，次の値を μ または σ^2 によって表せ．

（a） $E(2X)$　（b） $E(3X+1)$　（c） $E(X^2)$

（d） $E(2X^2+5X+7)$　（e） $V(3X)$　（f） $V(2X-3)$

（g） $V(3X+1)$　（h） $E(X^2+4X)$

解 $E(X)=\mu$, $V(X)\equiv\sigma^2=E(X^2)-\mu^2$ より

（a） $E(2X)=2E(X)=\boldsymbol{2\mu}$

（b） $E(3X+1)=3E(X)+1=\boldsymbol{3\mu+1}$

（c） $E(X^2)=\boldsymbol{\mu^2+\sigma^2}$

（d） $E(2X^2+5X+7)=2E(X^2)+5E(X)+7=2(\mu^2+\sigma^2)+5\mu+7$

$=\boldsymbol{2\mu^2+5\mu+2\sigma^2+7}$

（e） $V(3X)=9V(X)=\boldsymbol{9\sigma^2}$

(f)　$V(2X-3)=4V(X)=\boldsymbol{4\sigma^2}$

(g)　$V(3X+1)=9V(X)=\boldsymbol{9\sigma^2}$

(h)　$E(X^2+4X)=E(X^2)+4E(X)=\boldsymbol{\mu^2+\sigma^2+4\mu}$

例題 4　(独立な確率変数の1次結合の平均と分散)

X は平均 μ, 分散 σ^2 をもち, Y は平均 μ, 分散 $2\sigma^2$ をもつ. X と Y が独立のとき, 次の値を μ と σ^2 によって表せ.

(a)　$E(X-2)$　(b)　$V(X+Y)$　(c)　$V(X-Y)$

(d)　$V(2X+3Y)$　(e)　$V(X-3Y-5)$　(f)　$V(aX-bY)$

解　題意より $E(X)=E(Y)=\mu$, $V(X)=\sigma^2$, $V(Y)=2\sigma^2$ であることを用いる.

(a)　$E(X-2)=E(X)-2=\boldsymbol{\mu-2}$

(b)　$V(X+Y)=V(X)+V(Y)=\sigma^2+2\sigma^2=\boldsymbol{3\sigma^2}$

(c)　$V(X-Y)=V(X)+V(Y)=\boldsymbol{3\sigma^2}$

(d)　$V(2X+3Y)=4V(X)+9V(Y)=4\sigma^2+9\times 2\sigma^2=\boldsymbol{22\sigma^2}$

(e)　$V(X-3Y-5)=V(X)+9V(Y)=\sigma^2+9\times 2\sigma^2=\boldsymbol{19\sigma^2}$

(f)　$V(aX-bY)=a^2V(X)+b^2V(Y)=a^2\sigma^2+b^2\times 2\sigma^2=\boldsymbol{(a^2+2b^2)\sigma^2}$

例題 5　(ある離散型確率分布の平均と分散)

箱の中に4個の赤球と3個の青球がある. この箱から, 非復元抽出で青球が出るまで球のとり出しを続ける. X を青球が出るまでの球のとり出し回数とするとき, X の確率分布を求めよ. また, X の平均と分散を求めよ.

解　赤球をR, 青球をBで表す. X の確率分布は X がとる値とそれに対する確率を求めればよい.

x	1	2	3	4	5
事象系列	B	RB	RRB	RRRB	RRRRB
確率	$\frac{3}{7}$	$\frac{4}{7}\cdot\frac{3}{6}$	$\frac{4}{7}\cdot\frac{3}{6}\cdot\frac{3}{5}$	$\frac{4}{7}\cdot\frac{3}{6}\cdot\frac{2}{5}\cdot\frac{3}{4}$	$\frac{4}{7}\cdot\frac{3}{6}\cdot\frac{2}{5}\cdot\frac{1}{4}\cdot 1$
$P(X=x)$	$\frac{3}{7}$	$\frac{2}{7}$	$\frac{6}{35}$	$\frac{3}{35}$	$\frac{1}{35}$

よって，X の確率分布は

x	1	2	3	4	5
$P(X=x)$	$\frac{3}{7}$	$\frac{2}{7}$	$\frac{6}{35}$	$\frac{3}{35}$	$\frac{1}{35}$

ゆえに，

$$E(X)=1\times\frac{3}{7}+2\times\frac{2}{7}+3\times\frac{6}{35}+4\times\frac{3}{35}+5\times\frac{1}{35}=\mathbf{2}\ (\text{回})$$

$$V(X)=E(X^2)-2^2$$
$$=1^2\times\frac{3}{7}+2^2\times\frac{2}{7}+3^2\times\frac{6}{35}+4^2\times\frac{3}{35}+5^2\times\frac{1}{35}-2^2=\mathbf{1.2}\ (\text{回})$$

例題 6（**離散型確率分布の応用**）

バレーボールの試合は，どちらか一方のチームが先に 3 セットを勝てば試合は終了する．1 セットで A チームが B チームに勝つ確率を p，負ける確率を q とし，各セットの試合は独立で，p はゲームを通して一定とする．このとき，次を求めよ．

（a） 試合が 3 ゲームで終わる確率．

（b） 試合が 4 ゲームで終わる確率．

（c） 試合が終わるまでのセット数を X とするとき，X の確率分布．

（d） $p=\frac{1}{2}$ のとき，X の期待値と分散．

解（a） 試合が 3 ゲームで終わるのは，A が 3 連勝するか B が 3 連勝するかのいずれかであるから，求める確率は $\boldsymbol{p^3+q^3}$．

（b） 試合が 4 ゲームで終わるのは，A が第 4 セットに勝ち，通算 3 勝して終わるか，B が第 4 セットに勝ち，通算 3 勝して終わるかのいずれかである．

A が勝つ場合は，A の勝を○，負を×で表すと

第 1 セット	第 2 セット	第 3 セット	第 4 セット	確率
×	○	○	○	qp^3
○	×	○	○	qp^3
○	○	×	○	qp^3

の 3 通りで，その確率は $3qp^3$ となる．p と q を入れ替えれば B が勝つ確率 $3qp^3$ が得られるから，試合が 4 ゲームで終わる確率は

$$3qp^3+3pq^3=\mathbf{3pq(p^2+q^2)}$$

（c）　(b)と同様に考えて，試合が5ゲームで終わるのは第1セットから第4セットまでにAが2回勝ち，第5セットでAが勝つか，第4セットまでにBが2回勝ち，第5セットでBが勝つかのいずれかである．前者の確率は${}_4C_2p^2q^2\times p$で，後者の確率は${}_4C_2p^2q^2\times q$であるから，試合が5ゲームで終わる確率は

$$ {}_4C_2p^2q^2\times p+{}_4C_2p^2q^2\times q={}_4C_2p^2q^2=6p^2q^2 $$

試合は5ゲームまでに必ず終わるので，試合が終わるまでに要したセット数をXとすれば，Xの確率分布は

x	3	4	5
$P(X=x)$	p^3+q^3	$3pq(p^2+q^2)$	$6p^2q^2$

（d）　$p=\dfrac{1}{2}$のとき，Xの確率分布は

x	3	4	5
$P(X=x)$	$\frac{2}{8}$	$\frac{3}{8}$	$\frac{3}{8}$

よって，

$$ E(X)=3\times\frac{2}{8}+4\times\frac{3}{8}+5\times\frac{3}{8}=\mathbf{\frac{33}{8}} $$

$$ V(X)=3^2\times\frac{2}{8}+4^2\times\frac{3}{8}+5^2\times\frac{3}{8}-\left(\frac{33}{8}\right)^2=\mathbf{\frac{39}{64}} $$

例題　7　**（確率分布の応用）**

ある人は毎日車で会社に通い，会社の近くのA駐車場を利用している．A駐車場から会社までは歩いて8分である．彼の車がAに着いたとき，ここが空いておれば駐車するが，満杯のときはAから少し離れたB駐車場を利用することにしている．Bは十分なスペースをもつのでいつでも駐車できる．Bから会社までは歩いて15分，AからBまでは車で9分かかる．彼がA駐車場に着いたとき，そこが満杯である確率は常に1/4である．彼がA駐車場に着いてから彼の会社までX分かかるとき，Xの平均と標準偏差を求めよ．

解　この人の会社への行き方は，車をAに駐車して行くか，Bに駐車して行くかの2通りで，Aからの所要時間は8分，Aに駐車する確率は$\dfrac{3}{4}$，Bからの

所要時間は 9＋15＝24（分），B に駐車する確率は $\frac{1}{4}$ である．よって X の確率分布は

x	8	24
$P(X=x)$	$\frac{3}{4}$	$\frac{1}{4}$

ゆえに，

$$\mu=E(X)=8\times\frac{3}{4}+24\times\frac{1}{4}=\mathbf{12}\ (\text{分})$$

$$\sigma=\sqrt{V(X)}=\sqrt{8^2\times\frac{3}{4}+24^2\times\frac{1}{4}-12^2}=\mathbf{6.9}\ (\text{分})$$

例題 8　（結合確率分布）

次の 2 元表は 2 つの離散型確率変数の結合確率分布を示す．

（a）　X の周辺分布と Y の周辺分布を求めよ．

（b）　$E(X)$, $E(Y)$, $E(XY)$, $V(X)$, $V(Y)$ を求めよ．

（c）　$E(XY)=E(X)E(Y)$ は成り立つが，X と Y は独立ではないことを示せ．

（d）　$Z=X+Y$ の分散を求めよ．

x ＼ y	1	2	3
0	$\frac{3}{20}$	$\frac{1}{10}$	$\frac{3}{20}$
1	$\frac{1}{10}$	0	$\frac{1}{10}$
2	$\frac{3}{20}$	$\frac{1}{10}$	$\frac{3}{20}$

解　（a）　2 元表でそれぞれの周辺和を求めれば，

X の周辺分布

x	0	1	2
$P(X=x)$	$\frac{2}{5}$	$\frac{1}{5}$	$\frac{2}{5}$

Y の周辺分布

y	1	2	3
$P(Y=y)$	$\frac{2}{5}$	$\frac{1}{5}$	$\frac{2}{5}$

（b）　（a）より

$$E(X)=0\times\frac{2}{5}+1\times\frac{1}{5}+2\times\frac{2}{5}=\mathbf{1}$$

$$E(Y)=1\times\frac{2}{5}+2\times\frac{1}{5}+3\times\frac{2}{5}=\mathbf{2}$$

$$V(X)=0^2\times\frac{2}{5}+1^2\times\frac{1}{5}+2^2\times\frac{2}{5}-1^2=\mathbf{\frac{4}{5}}$$

$$V(Y)=1^2\times\frac{2}{5}+2^2\times\frac{1}{5}+3^2\times\frac{2}{5}-2^2=\mathbf{\frac{4}{5}}$$

$$E(XY)=1\times1\times\frac{1}{10}+1\times3\times\frac{1}{10}+2\times1\times\frac{3}{20}+2\times2\times\frac{1}{10}+2\times3\times\frac{3}{20}=\mathbf{2}$$

（c）（b）より

$$E(XY)=E(X)E(Y)$$

は成り立つ．しかし，たとえば

$$P(X=0,\ Y=1)=\frac{3}{20},\quad P(X=0)=\frac{2}{5},\quad P(Y=1)=\frac{2}{5}$$

より

$$P(X=0,\ Y=1)\neq P(X=0)P(Y=1)$$

であるから，X と Y は独立ではない．

（d）$Z=X+Y$ の確率分布は，2 元表よりの直接計算によって，

$z(=x+y)$	1	2	3	4	5
$P(Z=z)$	$\frac{3}{20}$	$\frac{1}{5}$	$\frac{3}{10}$	$\frac{1}{5}$	$\frac{3}{20}$

であるから

$$E(Z)=\mathbf{3} \qquad \text{（分布の対称性より）}$$

$$V(Z)=1^2\times\frac{3}{20}+2^2\times\frac{1}{5}+3^2\times\frac{3}{10}+4^2\times\frac{1}{5}+5^2\times\frac{3}{20}-3^2=\mathbf{1.6}$$

例題 9 （確率密度関数）

関数

$$f(x)=\frac{5}{8}(1-x^4) \qquad (-1\leqq x\leqq 1)$$

は連続型確率変数の密度関数であることを示し，

（a）$P\left(X>\frac{1}{2}\right)$

（b）$P\left(X^2<\frac{1}{4}\right)$

を求めよ．

解　ある関数 $f(x)$ が密度関数であることを示すには

（i）　すべての x に対して $f(x) \geqq 0$

（ii）　$\int_{-\infty}^{\infty} f(x)\,dx = 1$

の2条件が成り立つことをいえばよい.

明らかに, $\frac{5}{8}(1-x^4) \geqq 0$ $(-1 \leqq x \leqq 1)$ だから,（i）は成立.

$$\int_{-1}^{1} \frac{5}{8}(1-x^4)\,dx = \frac{5}{8}\left[x - \frac{x^5}{5}\right]_{-1}^{1} = 1$$

より (ii) も成立.

よって, 与えられた関数は密度関数である.

（a）　$P\left(X > \frac{1}{2}\right) = \int_{\frac{1}{2}}^{1} \frac{5}{8}(1-x^4)\,dx = \frac{5}{8}\left[x - \frac{x^5}{5}\right]_{\frac{1}{2}}^{1} = \mathbf{\frac{49}{256}}$

（b）　$P\left(X^2 < \frac{1}{4}\right) = P\left(-\frac{1}{2} < X < \frac{1}{2}\right)$

$$= \int_{-\frac{1}{2}}^{\frac{1}{2}} \frac{5}{8}(1-x^4)\,dx = \frac{5}{8}\left[x - \frac{x^5}{5}\right]_{-\frac{1}{2}}^{\frac{1}{2}} = \mathbf{\frac{79}{128}}$$

例題 10　（密度関数の平均・分散・モード）

X の確率密度関数が

$$f(x) = \begin{cases} c(1-x)(x-2) & (1 \leqq x \leqq 2) \\ 0 & (\text{その他}) \end{cases}$$

のとき,

（a）　c の値,

（b）　X の平均,

（c）　X の分散,

（d）　X のモード

を求めよ.

解　（a）　$\int_1^2 c(1-x)(x-2)\,dx = 1$ より

$$c\int_1^2 (-2+3x-x^2)\,dx = c\left[-2x + \frac{3}{2}x^2 - \frac{x^3}{3}\right]_1^2 = 1 \quad \Rightarrow \quad c = \mathbf{6}$$

（b）　$E(X) = 6\int_1^2 x(1-x)(x-2)\,dx$

$$=6\int_1^2(-2x+3x^2-x^3)\,dx$$

$$=6\left[-x^2+x^3-\frac{x^4}{4}\right]_1^2=\mathbf{\frac{3}{2}}$$

(c)　$V(X)=E(X^2)-\{E(X)\}^2$

$$=6\int_1^2 x^2(1-x)(x-2)\,dx-\frac{9}{4}$$

$$=6\left[-\frac{2}{3}x^3+\frac{3}{4}x^4-\frac{x^5}{5}\right]_1^2-\frac{9}{4}$$

$$=\frac{23}{10}-\frac{9}{4}=\mathbf{\frac{1}{20}}$$

(d)　$f(x)=6(1-x)(x-2)$

$f'(x)=18-12x=0 \ \Rightarrow\ x=\frac{3}{2}$. よって，モードは $\mathbf{\frac{3}{2}}$.

例題 11　(密度関数のメジアン)

X の確率密度関数が

$$f(x)=\begin{cases}\dfrac{a}{x^3} & (1\leqq x\leqq 2)\\ 0 & (\text{その他})\end{cases}$$

のとき，

(a)　a の値，

(b)　$P(X\leqq y)=2P(X\geqq y)$ を満たす y，

(c)　X のメジアン，

を求めよ．

解　(a)　$\int_1^2 \frac{a}{x^3}\,dx=1$ より

$$a\left[-\frac{1}{2x^2}\right]_1^2=1 \ \Rightarrow\ a=\mathbf{\frac{8}{3}}$$

(b)　$P(X\leqq y)=\frac{8}{3}\int_1^y\frac{dx}{x^3}=\frac{8}{3}\left[-\frac{1}{2x^2}\right]_1^y=\frac{8}{3}\left(\frac{1}{2}-\frac{1}{2y^2}\right)$

$$2P(X\geqq y)=\frac{16}{3}\int_y^2\frac{dx}{x^3}=\frac{16}{3}\left(\frac{1}{2y^2}-\frac{1}{8}\right)$$

$$\frac{8}{3}\left(\frac{1}{2}-\frac{1}{2y^2}\right)=\frac{16}{3}\left(\frac{1}{2y^2}-\frac{1}{8}\right)$$

より，$y=\mathbf{\sqrt{2}}$.

（c）　メジアンを M とすると

$$\frac{8}{3}\int_1^M \frac{dx}{x^3}=\frac{1}{2}$$

$$\frac{8}{3}\left(\frac{1}{2}-\frac{1}{2M^2}\right)=\frac{1}{2}$$

$$M^2=\frac{8}{5} \quad \Rightarrow \quad M=\boldsymbol{\frac{2\sqrt{10}}{5}}$$

例題 12　**（密度関数と分布関数）**

X の確率密度関数が

$$f(x)=\begin{cases} ax(2-x) & (0 \leqq x \leqq 2) \\ 0 & (\text{その他}) \end{cases}$$

のとき，

（a）　a の値を求めよ．

（b）　$P(0 \leqq X \leqq 1)$ を求めよ．

（c）　X の平均は 1 で，分散は $\frac{1}{5}$ であることを示せ．

（d）　X の累積分布関数を求めよ．

解　（a）

$$\int_0^2 ax(2-x)\,dx=1$$

$$a\int_0^2 (2x-x^2)\,dx=1$$

$$a\left[x^2-\frac{x^3}{3}\right]_0^2=1 \quad \Rightarrow \quad a=\boldsymbol{\frac{3}{4}}$$

（b）　$P(0 \leqq X \leqq 1)=\frac{3}{4}\int_0^1 x(2-x)\,dx$

$$=\frac{3}{4}\left[x^2-\frac{x^3}{3}\right]_0^1=\boldsymbol{\frac{1}{2}}$$

（c）　$E(X)=\int_0^2 x\cdot\frac{3}{4}x(2-x)\,dx$

$$=\frac{3}{4}\left[\frac{2x^3}{3}-\frac{x^4}{4}\right]_0^2=\mathbf{1}$$

$V(X)=E(X^2)-\{E(X)\}^2$

$$=\frac{3}{4}\int_0^2 x^3(2-x)\,dx-1^2$$

$$=\frac{3}{4}\left[\frac{x^4}{2}-\frac{x^5}{5}\right]_0^2-1=\boldsymbol{\frac{1}{5}}$$

（d）　$P(X<0)=0$ より，$x<0$ に対しては $F(x)=0$.

$P(X>2)=1$ より，$x>2$ に対しては $F(x)=1$.

$0 \leqq x \leqq 2$ の範囲では，

$$F(x)=\int_0^x \frac{3}{4}x(2-x)\,dx$$
$$=\frac{3}{4}\left(x^2-\frac{x^3}{3}\right)+c$$

$x=0$ のとき，$F(x)=0$ だから，$c=0$.

よって，

$$F(x)=\begin{cases} 0 & (x<0) \\ \dfrac{3}{4}\left(x^2-\dfrac{x^3}{3}\right) & (0 \leqq x \leqq 2) \\ 1 & (x>2) \end{cases}$$

例題 13　**（密度関数の応用問題）**

列車がR駅に到着するときの誤差 X（分）の確率密度関数が

$$f(x)=\begin{cases} c(16-x^2) & (-4 \leqq x \leqq 4) \\ 0 & (\text{その他}) \end{cases}$$

で与えられるとき，c の値を定めよ．この列車が定刻より

（a）　少なくとも2分遅れる，

（b）　少なくとも1分早く着く，

（c）　1分から3分遅れる

確率を求めよ．

解　$1=\int_{-4}^{4} c(16-x^2)\,dx=c\left[16x-\dfrac{x^3}{3}\right]_{-4}^{4}=\dfrac{256}{3}c \quad \Rightarrow \quad c=\mathbf{\dfrac{3}{256}}$

列車が駅に到着するときの誤差を X とすると，題意より

（a）　$P(X>2)=\int_2^4 \dfrac{3}{256}(16-x^2)\,dx=\dfrac{3}{256}\left[16x-\dfrac{x^3}{3}\right]_2^4=\mathbf{\dfrac{5}{32}}$

（b）　$P(X<-1)=\int_{-4}^{-1} \dfrac{3}{256}(16-x^2)\,dx=\dfrac{3}{256}\left[16x-\dfrac{x^3}{3}\right]_{-1}^{-4}=\mathbf{\dfrac{81}{256}}$

（c）　$P(1<X<3)=\int_1^3 \dfrac{3}{256}(16-x^2)\,dx=\dfrac{3}{256}\left[16x-\dfrac{x^3}{3}\right]_1^3=\mathbf{\dfrac{35}{128}}$

例題 14　（分布関数と密度関数）

X の累積分布関数が

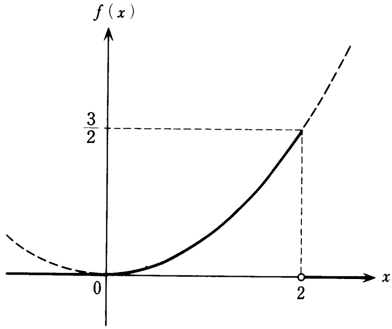
$$F(x)=\begin{cases} 0 & (x<0) \\ kx^3 & (0<x\leqq 2) \\ 1 & (x>2) \end{cases}$$

のとき,

（a）　X の確率密度関数 $f(x)$,

（b）　X の平均と分散

を求め, $f(x)$ のグラフを示せ.



解　（a）　$f(x)=\dfrac{dF(x)}{dx}=3kx^2$

$F(2)=1$ より

$$8k=1 \quad\Rightarrow\quad k=\frac{1}{8}$$

よって,

$$f(x)=\begin{cases} \dfrac{3}{8}x^2 & (0\leqq x\leqq 2) \\ 0 & (\text{その他}) \end{cases}$$

（b）　$E(X)=\displaystyle\int_0^2 x\cdot\frac{3}{8}x^2\,dx=\frac{3}{8}\left[\frac{x^4}{4}\right]_0^2=\boldsymbol{\frac{3}{2}}$

$$V(X)=\int_0^2 x^2\frac{3}{8}x^2\,dx-\left(\frac{3}{2}\right)^2=\frac{3}{8}\left[\frac{x^5}{5}\right]_0^2-\frac{9}{4}=\boldsymbol{\frac{3}{20}}$$

$f(x)$ のグラフは右上に図示したとおり.

例題 15

あるガソリンスタンドは毎週月曜日の朝, ガソリンの補給を受ける. このスタンドの週当たりガソリン販売量を X（1000 リットル単位）とする. 過去の経験から, X の確率密度関数は

$$f(x)=\frac{3}{125}(5-x)^2 \qquad (0\leqq x\leqq 5)$$

であることが知られている.

（a）　このスタンドのある週の販売量が 2000 リットル未満である確率を求めよ.

（b）　このスタンドのガソリンタンクの容量は 4000 リットルであるとき，ある週に，このスタンドがガソリンの需要を満たせない確率を求めよ．

解　（a）　週販売量 X は 1000 リットル単位で測られているから

$$P(X<2)=\int_0^2 \frac{3}{125}(5-x)^2\,dx$$
$$=\frac{3}{125}\int_3^5 y^2\,dy \qquad (5-x=y \text{ とおく})$$
$$=\frac{3}{125}\left[\frac{y^3}{3}\right]_3^5=\frac{98}{125}=\mathbf{0.784}$$

（b）　需要を満たせないのは，X が 4000 リットルを超えるときであるから

$$P(X>4)=\int_4^5 \frac{3}{125}(5-x)^2\,dx \qquad (5-x=y \text{ とおく})$$
$$=\frac{3}{125}\left[\frac{y^3}{3}\right]_0^1=\frac{1}{125}=\mathbf{0.008}$$

3 章の問題

3.1　次の各確率分布の平均と分散を求めよ．

（a）

x	0	2	5	9
$P(X=x)$	0.4	0.1	0.2	0.3

（b）

x	-2	-1	0	1	2
$P(X=x)$	0.1	0.2	0.4	0.2	0.1

（c）

x	-2	2	4	5
$P(X=x)$	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{3}{8}$	$\frac{1}{4}$

3.2　X の確率分布が

x	0	1	4	6	9
$P(X=x)$	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{2}$	$\frac{1}{6}$	$\frac{1}{12}$

のとき，この確率分布のグラフを示せ．$E(X)$，$V(X)$ を求めよ．また，$Y=5X+2$ のとき，$E(Y)$，$V(Y)$ を求めよ．

3.3 X の確率分布が

x	0	1	2	3	4
$P(X=x)$	0.1	0.1	0.3	0.3	0.2

のとき，

(a) X の期待値と分散を求めよ．

(b) X^2 の期待値と分散を求めよ．

3.4 X が確率分布

x	0	1	2	3
$P(X=x)$	p	p^2	$2p^2$	p

をもつとき，X の平均と分散を求めよ．

3.5 X の平均が 2，分散が 5 のとき．

(a) $X-1$　(b) $3X$　(c) $2X+1$　(d) $\frac{1}{3}(X+3)$

(e) $5-3X$

の平均と分散を求めよ．

3.6 数 1, 2, 3, 4, 5, 6 が記入された 6 個の球の入った箱がある．この箱から，非復元抽出で 2 個の球をとり出す．とり出された球の数の和を X，積を Y とするとき，X と Y の確率分布をそれぞれ求めよ．

3.7 4 人で競うトランプゲームで，ある 1 人に配られた 13 枚の札の中の“エース”の数を X とするとき，確率変数 X の確率分布を求めよ．

3.8 次の式で正しくないのはどれか．その理由は．

(a) $E(X+X)=E(X)+E(X)=2E(X)$

(b) $V(X+X)=V(X)+V(X)=2V(X)$

(c) $V(X+X)=V(2X)=4V(X)$

3.9 サイコロを 2 回投げ，最初に出た目を X，2 回目に出た目を Y とするとき，

(a)　$Z=|X-Y|$

(b)　$W=\min(X, Y)$

の確率分布をそれぞれ求めよ．また，Z と W の平均と分散をそれぞれ求めよ．

3.10　X と Y の結合確率分布が

x \ y	0	1	2
0	$\frac{1}{60}$	$\frac{1}{30}$	$\frac{1}{20}$
1	$\frac{1}{30}$	$\frac{1}{15}$	$\frac{1}{10}$
2	$\frac{1}{20}$	$\frac{1}{10}$	$\frac{3}{20}$
3	$\frac{1}{15}$	$\frac{2}{15}$	$\frac{1}{15}$

のとき，

(a)　X と Y の周辺分布をそれぞれ求めよ．

(b)　X と Y は独立かどうか．

(c)　$2X-Y$ の平均と分散を求めよ．

3.11　(X, Y) の結合確率分布が

x \ y	-1	0	1	2
1	$\frac{1}{8}$	$\frac{2}{8}$	0	0
2	$\frac{1}{8}$	0	$\frac{3}{8}$	$\frac{1}{8}$

のとき，次を求めよ．

(a)　X の周辺分布と Y の周辺分布．

(b)　$E(Y)$ と $V(Y)$．

(c)　X と Y は独立かどうか．

(d)　$E(X+Y)$．

(e)　$E(XY)$．

3.12　赤球2個，青球1個，白球5個を含む袋から非復元抽出で4個の球をとり出すとき，赤球1個と白球3個がとり出される確率は $\frac{2}{7}$ になることを

示せ．この袋から非復元抽出で 4 個の球をとり出すとき，得られた赤球の数を X，青球の数を Y とする．そのとき，X と Y の結合確率分布を与える 2 元表を示せ．この表から X と Y は独立でないことを示せ．$Z=X+Y$ の確率分布を導き，その平均と分散を求めよ．

3.13　ある自動車セールスマンの基本給は月 12 万円で，新車を 1 台売るごとに 3 万円の歩合がもらえる．セールスマンの月間新車販売台数は確率変数で，その平均は 1.85 台，標準偏差は 1.24 台である．このセールスマンの月間収入の平均と標準偏差を求めよ．

3.14　X の密度関数が

$$f(x)=\begin{cases} ax^2 & (0 \leqq x \leqq 1) \\ 0 & (\text{その他}) \end{cases}$$

のとき，次を求めよ．

（a）　a の値．

（b）　$P\left(X>\dfrac{1}{2}\right)$.

（c）　X の平均と分散．

（d）　X の分布関数．

3.15　X の確率密度関数が

$$f(x)=\begin{cases} a+bx+cx^2 & (0 \leqq x \leqq 1) \\ \quad 0 & (\text{その他}) \end{cases}$$

で，その平均は $\dfrac{1}{2}$，分散は $\dfrac{1}{20}$ のとき，

（a）　a, b, c の値を求めよ．

（b）　この分布のモードとメジアンを求めよ．

3.16　X の密度関数が

（a）　$f(x)=\dfrac{3}{10}x(3-x) \qquad (1 \leqq x \leqq 3)$

（b）　$f(x)=\dfrac{3}{2}(1-x^2) \qquad (0 \leqq x \leqq 1)$

（c）　$f(x)=\begin{cases} 1+x & (-1 \leqq x \leqq 0) \\ 1-x & (0<x<1) \end{cases}$

で与えられるとき，各分布の平均と分散を求めよ．

3.17　次の密度関数から分布関数を求めよ.

（a）　$f(x)=\dfrac{3}{4}(1-x^2)$　　$(-1\leqq x\leqq 1)$

（b）　$f(x)=3(1-x)^2$　　$(0\leqq x\leqq 1)$

（c）　$f(x)=\dfrac{3}{32}x(4-x)$　　$(0\leqq x\leqq 4)$

3.18　シリコンチップの寿命 X の確率密度関数が

$$f(x)=\begin{cases}\dfrac{200}{x^2} & (x\geqq 200)\\ 0 & (\text{その他})\end{cases}$$

で与えられるとき，無作為に選ばれた2個のチップが両方とも500時間以内に取り替えねばならない確率を求めよ.

3.19　ある人が半径4 cmの円形の標的に向けてライフルを発射する．標的は中心の半径がそれぞれ1 cm，2 cm，3 cmの同心円からなる．標的の中心から着弾点までの距離 X は確率変数で，その確率密度関数が

$$f(x)=0.03(x^2+3)\qquad (0\leqq x\leqq 4)$$

であるとき，

（a）　弾丸が小さい円に当たる確率を求めよ.

（b）　弾丸が小さい円に当たると5点，この円と中間の円の間に当たると3点，中間の円と大きい円の間に当たると1点，大きい円の外にはずれたら0点が与えられる．この場合，

（i）　確率が最も大きいのは何点のときか.

（ii）　1発当たりの得点の平均を求めよ.

3.20　X の確率密度関数が

$$f(x)=\frac{1}{2}e^{-|x|}\qquad (-\infty<x<\infty)$$

のとき，$f(x)$ のグラフを図示し，X の平均と分散を求めよ.

4

2 項分布とポアソン分布

4-1　2 項 分 布

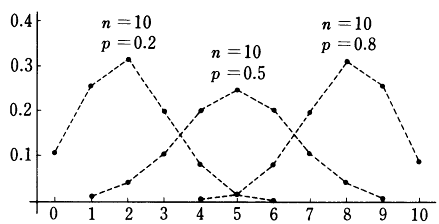
2 項分布　1 回の試行である事象の起こる確率を p とする．n 回の独立試行でこの事象の起こる回数 X の確率分布は

$$P(X=x)={}_n\mathrm{C}_x p^x q^{n-x} \qquad (x=0, 1, 2, \cdots, n\,;\, q=1-p)$$

で与えられる．この分布を **2 項分布**といい，記号 $B(n, p)$ で表す．ここで，${}_n\mathrm{C}_x$ は 2 項係数で，${}_n\mathrm{C}_x=\dfrac{n!}{x!(n-x)!}$ である．X が 2 項分布 $B(n, p)$ に従うことを

$$X \sim B(n, p)$$

と書く．2 項分布は表の形で示すと次のようになる．



2 項分布

x	0	1	2	$\cdots$	x	$\cdots$	n	計
$P(X=x)$	q^n	npq^{n-1}	${}_n\mathrm{C}_2 p^2 q^{n-2}$	$\cdots$	${}_n\mathrm{C}_x p^x q^{n-x}$	$\cdots$	p^n	1

表のなかの各確率は 2 項展開

$$(q+p)^n=q^n+{}_n\mathrm{C}_1 pq^{n-1}+{}_n\mathrm{C}_2 pq^{n-2}+\cdots+p^n$$

の各項である．

2 項分布の平均と分散

$$\mu=np$$

$$\sigma^2=npq$$

2 項分布のモード

（i）　$(n+1)p$ が整数ならば，モードは

$$(n+1)p \text{ と } (n+1)p-1 \text{ の 2 つ.}$$

（ii）　$(n+1)p$ が整数でないならば，モードは

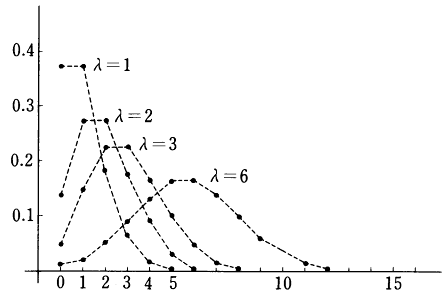
$$(n+1)p \text{ を超えない最大の整数値.}$$

4-2　ポアソン分布

ポアソン分布　　X の確率分布が

$$P(X=x)=\frac{\lambda^x e^{-\lambda}}{x!} \qquad (x=0, 1, 2, \cdots)$$

で与えられるとき，この分布を母数 λ の**ポアソン分布**という．n が十分大きく，p は非常に小さく，$np=\lambda$ のときの 2 項分布は母数 λ のポアソン分布で近似される．ポアソン分布はまた，時間または空間内でランダムに起こる現象の単位時間または単位体積内での生起数に対する理論分布として使われる．たとえば，ある物質が単位時間内に放出する放射性粒子の数や，電話交換機が単位時間に受ける電話の呼出し数などはポアソン分布に従う．



ポアソン分布

ポアソン分布の平均と分散

$$\mu=\lambda$$

$$\sigma^2=\lambda$$

2 項分布のポアソン分布による近似　　2 項分布 $B(n, p)$ がポアソン分布で近似できるための一般的条件は

$$n>50, \qquad np \leqq 5$$

4-3　その他の重要な離散型確率分布

離散型一様分布　　X の確率分布が

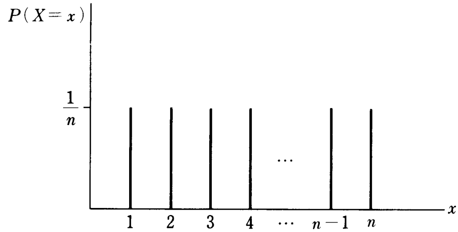
$$P(X=x)=\begin{cases} \dfrac{1}{n} & (x=1, 2, \cdots, n) \\ 0 & (\text{その他}) \end{cases}$$

で与えられるとき，この分布を**離散型一様分布**という．

離散型一様分布の平均と分散

$$\mu=\frac{n+1}{2}$$

$$\sigma^2=\frac{n^2-1}{12}$$



一様分布

幾何分布　X の確率分布が

$$P(X=x)=p(1-p)^{x-1} \qquad (x=1, 2, \cdots; 0 \leqq p<1)$$

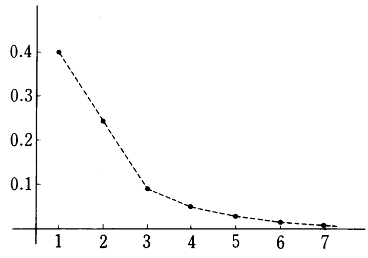
で与えられるとき，

この分布を**幾何分布**という．

幾何分布の平均と分散

$$\mu=\frac{1}{p}$$

$$\sigma^2=\frac{1-p}{p^2}$$



$p=0.4$ の幾何分布

例　題

例題 1　**(2 項分布の応用)**

ある機械が作る部品の 20 % は不良品である．この機械で作られた部品を無作為に 8 個とるとき，その中に

(a)　2 個の不良品が含まれる，

(b)　高々1 個の不良品が含まれる，

(c)　少なくとも 3 個の不良品が含まれる

確率を求めよ．

解　8 個の部品のなかの不良品の数を X とすると，X は $n=8$, $p=0.2$ の 2 項分布

$$P(X=x)={}_8C_x(0.2)^x(0.8)^{8-x} \qquad (x=0, 1, 2, \cdots, 8)$$

に従う．よって

(a)　$P(X=2)\equiv P(2)={}_8C_2(0.2)^2(0.8)^6=\mathbf{0.294}$

(b)　$P(X\leqq 1)=P(0)+P(1)$

$$=(0.8)^8+8(0.2)(0.8)^7=\mathbf{0.382}$$

(c)　$P(X\geqq 3)=1-p(0)-P(1)-P(2)=\mathbf{0.324}$

例題 2　(2項分布の平均と分散)

X は2項分布 $B(n, 0.8)$ に従い，

$$P(X=4)=5P(X=3)$$

である．この2項分布の平均と分散を求めよ．

解　$X \sim B(n, 0.8)$ より

$$P(X=x)={}_n\mathrm{C}_x(0.8)^x(0.2)^{n-x} \qquad (x=0, 1, 2, \cdots, n)$$

それゆえ，

$$P(X=4)={}_n\mathrm{C}_4(0.8)^4(0.2)^{n-4}$$
$$P(X=3)={}_n\mathrm{C}_3(0.8)^3(0.2)^{n-3}$$

与えられた関係より

$$5=\frac{P(X=4)}{P(X=3)}=\frac{\dfrac{n!}{4!(n-4)!}}{\dfrac{n!}{3!(n-3)!}}\times\frac{0.8}{0.2}=n-3 \qquad \therefore\ n=8$$

よって，2項分布の平均と分散の公式から

$$E(X)=np=8\times 0.8=\mathbf{6.4}$$
$$V(X)=npq=8\times 0.8\times 0.2=\mathbf{1.28}$$

例題 3　(2項分布の平均と分散)

2項分布 $B(2, p)$ と $B(3, p)$ の平均と分散を，公式からでなく，直接計算によって導け．

解　2項分布 $B(2, p)$ の各確率は

$$(q+p)^2=q^2+2pq+p^2$$

の各項であるから，$B(2, p)$ は

x	0	1	2
$P(X=x)$	q^2	$2pq$	p^2

と書ける．ゆえに，この分布の平均と分散は

$$E(X)=0\times q^2+1\times 2pq+2\times p^2=2p(q+p)=2p \qquad (\because\ p+q=1)$$

$$\begin{aligned}V(X)&=E(X^2)-\{E(X)\}^2\\&=0^2\times q^2+1^2\times 2pq+2^2\times p^2-(2p)^2=2pq\end{aligned}$$

同様に，2 項分布 $B(3, p)$ の各確率は

$$(q+p)^3=q^3+3pq^2+3p^2q+p^3$$

の各項であるから，$B(3, p)$ は

x	0	1	2	3
$P(X=x)$	q^3	$3pq^2$	$3p^2q$	p^3

と書ける．よって，

$$\begin{aligned} E(X)&=0\times q^3+1\times 3pq+2\times 3p^2q+3\times p^3 \\ &=3p(q^2+2pq+p^2)=3p(q+p)^2=\mathbf{3p} \\ V(X)&=0^2\times q^3+1^2\times 3pq+2^2\times 3p^2q+3^2\times p^3-(3p)^2 \\ &=3pq(q+4p)-9p^2(1-p) \\ &=3pq(q+p+3p)-9p^2q=\mathbf{3pq} \end{aligned}$$

［注］　同様な計算によって，2 項分布 $B(4, p)$ の平均と分散は，$4p$ と $4pq$ となることも簡単に示される．これらより，一般の 2 項分布 $B(n, p)$ の平均と分散は np と npq となることが類推される．

例題 4　**（2 項分布の応用）**

ある商品のセールスマンの基本給は月 10 万円で，商品 1 個売るごとに歩合 5000 円がもらえる．セールスマンは毎月 100 軒の家を訪門し，彼の訪れた家がこの商品を買う確率は 0.2 であるとする．1 人のセールスマンの月間販売量の平均と分散を求めよ．また，セールスマンの月間収入の平均と分散を求めよ．

解　セールスマンの毎月の販売量を X とすると，X は $n=100$, $p=0.2$ の 2 項分布に従う確率変数である．よって，

$$E(X)=np=100\times 0.2=20$$

$$V(X)=npq=100\times 0.2\times 0.8=16$$

セールスマンの毎月の収入を Y とすると

$$Y=10+0.5X$$

ゆえに，

$$E(Y)=10+0.5E(X)=10+0.5\times 20=\mathbf{20}\ \text{(万円)}$$

$$V(Y)=(0.5)^2\,V(X)=0.25\times 16=\mathbf{4}\ \text{(万円)}$$

例題 5 （2項分布の応用）

ある集団には左ききの人が10％いるといわれている．この集団から n 人を選んで左ききか否かを調べ，その中の少なくとも1人が左ききである確率を0.95以上にするには，少なくとも何人を選ばねばならないか．

解　n 人中の左ききの数を X とすると

$$X \sim B(n, 0.10)$$

与えられた条件から

$$P(X \geqq 1) = 1 - (0.9)^n \geqq 0.95$$

よって，

$$(0.9)^n \leqq 0.05$$

$$n \geqq \frac{\log 0.05}{\log 0.9} = 28.4$$

ゆえに，**29人以上**.

例題 6 （抜取検査）

ある検査法は，非常に大きい仕切りから無作為に8個の標本をとり，その中の不良品の数が2個以上のときは仕切りを不合格とし，不良品の数が0のときは合格とする．もし不良品の数が1個のときは，仕切りからさらに5個の標本をとり，その中に不良品がなければ合格，少なくとも1個あるときは不合格とする．

仕切り不良率が10％のとき，次の確率を求めよ．

（a）　第1回の標本で仕切りが合格となる．

（b）　第2回の標本で仕切りが合格となる．

（c）　この検査法で仕切りが合格となる．

解　X を第1回の標本での不良品の数とし，Y を第2回の標本での不良品の数とする．仕切りは十分大きいと仮定しているから，

X は近似的に2項分布 $B(8, 0.1)$ に従い，

Y は近似的に2項分布 $B(5, 0.1)$ に従う．

すなわち，

$$P(X=x) = {}_8\mathrm{C}_x (0.1)^x (0.9)^{8-x} \qquad (x = 0, 1, 2, \cdots, 8)$$

$$P(Y=y) = {}_5\mathrm{C}_y (0.1)^y (0.9)^{5-y} \qquad (y = 0, 1, 2, \cdots, 5)$$

よって

（a）　P（第1回の標本で仕切りが合格）

$=P(X=0)=(0.9)^8=\mathbf{0.430}$

（b）　P（第2回の標本で仕切りが合格）

$=P(X=1)P(Y=0)$　　（X と Y は独立であるから）

$=8(0.1)(0.9)^7\times(0.9)^5=\mathbf{0.226}$

（c）　P（仕切りが合格）

$=P$(第1回で仕切りが合格)$+P$(第2回で仕切りが合格)

$=0.430+0.226=\mathbf{0.656}$

例題 7　（ポアソン分布の応用）

りんごを250個ずつ箱に詰める．箱詰めされたりんごは平均して0.8％が腐るという．箱詰めりんご1箱をあけたとき，腐ったりんごが3個以上見出される確率を求めよ．

解　1箱中の腐ったりんごの数を X とすると，X は $n=250$, $p=0.008$ の2項分布に従う．この2項分布は n が50より大きく，p は非常に小さく，$\lambda=np=250\times0.008=2\leqq5$ であるから，$\lambda=2$ のポアソン分布

$$P(X=r)=\frac{2^r e^{-2}}{r!},\quad (r=0, 1, 2, \cdots)$$

で近似できる．よって，求める確率は

$$\begin{aligned} P(X\geqq3)&=1-P(0)-P(1)-P(2)\\ &=1-e^{-2}-2e^{-2}-2e^{-2}\\ &=1-5e^{-2}=\mathbf{0.32} \end{aligned}$$

例題 8　（ポアソン分布の応用）

A駅の売店でのある月刊雑誌の販売数は平均2冊のポアソン分布に従う．売店では毎月この雑誌を3冊仕入れている．

（a）　この店がある月，客の需要を満たせなくなる確率を求めよ．

（b）　この店でこの雑誌が1月当たりに売れる平均販売数を求めよ．

（c）　この店が毎月の雑誌の需要を満たす確率を少なくとも0.95にするには毎月最低何冊を仕入れねばならないか．

解　この店の毎月の雑誌販売数を X とすると

$$P(X=x)=\frac{2^x e^{-2}}{x!} \qquad (x=0, 1, 2, \cdots)$$

（a）　客の需要が満たせないのは，$X \geqq 4$ のときだから

$$\begin{aligned} P(X \geqq 4) &= 1-P(X \leqq 3) \\ &= 1-P(0)-P(1)-P(2)-P(3) \\ &= 1-e^{-2}-2e^{-2}-2e^{-2}-\frac{4}{3}e^{-2} \\ &= 1-\frac{19}{3}e^{-2}=\mathbf{0.14} \end{aligned}$$

（b）　このポアソン分布の平均は 2 だから，**2 冊**.

（c）　求める冊数を n とすると

$$P(X \leqq n) \geqq 0.95$$

$$e^{-2}\sum_{x=0}^{n}\frac{2^x}{x!} \geqq 0.95$$

$$\sum_{x=0}^{n}\frac{2^x}{x!} \geqq 7.01$$

ここで，$f_n=\sum_{x=0}^{n}\frac{2^x}{x!}$ とおくと，

$$f_3=6.3, \qquad f_4=7, \qquad f_5=7.2$$

よって，$n \geqq 5$. 最低 **5 冊**仕入れねばならない.

例題 9　**（ポアソン分布の応用）**

ある溶液は 1 ml 当たり平均 3 個のバクテリアを含む. この溶液 1 ml 中のバクテリアの数はポアソン分布に従うと仮定して，次の確率を求めよ.

（a）　1 ml の標本をとるとき，そのなかに 5 個以上のバクテリアが含まれる.

（b）　1 ml ずつ 2 個の標本をとるとき，どちらもそのなかにバクテリアを含まない.

（c）　1 ml ずつ 3 個の標本をとるとき，3 個のうち 2 個が少なくとも 1 個のバクテリアを含む.

解　この溶液 1 ml 中のバクテリアの数を X とすると，X は $\lambda=3$ のポアソン分布に従うから

$$P(X=x)=\frac{3^x e^{-3}}{x!} \qquad (x=0, 1, 2, \cdots)$$

よって，

（a）$P(X\geqq 5)=1-P(0)-P(1)-P(2)-P(3)-P(4)$

$$=1-e^{-3}-3e^{-3}-\frac{9}{2}e^{-3}-\frac{9}{2}e^{-3}-\frac{27}{8}e^{-3}$$

$$=1-\frac{131}{8}e^{-3}\fallingdotseq \mathbf{0.185}$$

（b）$[P(0)]^2=e^{-6}\fallingdotseq \mathbf{0.002}$

（c）1 ml の標本が少なくとも 1 個のバクテリアを含む確率は $P(X\geqq 1)$ で

$$P(X\geqq 1)=1-P(0)=1-e^{-3}\fallingdotseq 0.95$$

よって，3 個の標本のうち 2 個が少なくとも 1 個のバクテリアを含む確率は

$${}_3C_2(0.95)^2(0.05)\fallingdotseq \mathbf{0.135}$$

例題 10　**(2 項分布のポアソン近似)**

大都市では平均 80 人に 1 人が α 型の血液をもつという．

（a）無作為に 200 人の血液提供者を選ぶとき，その中に α 型の血液の人が少なくとも 4 人含まれる確率を求めよ．

（b）α 型の血液提供者をその中に少なくとも 1 人含む確率を 0.9 以上にするには何人を選ばねばならないか．

解　200 人中血液が α 型の人の数を X とすると

$$X \sim B\left(200, \frac{1}{80}\right)$$

この 2 項分布は，$n=200$ が大きく，$p=\frac{1}{80}$ は十分小さく，$\lambda=np=200\times\frac{1}{80}=2.5$ は 5 より小さいから，$\lambda=2.5$ のポアソン分布

$$P(X=x)=\frac{(2.5)^x e^{-2.5}}{x!} \qquad (x=0, 1, 2, \cdots)$$

で近似できる．よって，

（a）$P(X\geqq 4)=1-P(0)-P(1)-P(2)-P(3)$

$$=1-e^{-2.5}-2.5e^{-2.5}-\frac{(2.5)^2 e^{-2.5}}{2!}-\frac{(2.5)^3 e^{-2.5}}{3!}$$

$$\fallingdotseq \mathbf{0.24}$$

(b) $X \sim B\left(n, \dfrac{1}{80}\right)$. 題意より

$$P(X \geqq 1) \geqq 0.9$$
$$1-P(X=0) \geqq 0.9$$
$$\left(\frac{79}{80}\right)^n \leqq 0.1$$
$$n \geqq \frac{\log 0.1}{\log 0.9875} \fallingdotseq 183.1$$

よって, **184 人以上**.

例題 11 (ポアソン分布の当てはめ)

次の表は, 高速道路のある地点で観測した車の交通量の度数分布である.

車の数(10 秒間ごとの)	0	1	2	3	4	計
観測度数	68	81	38	9	4	200

データから平均と分散を求めよ. この分布にポアソン分布を当てはめたときの理論度数を求めよ.

解 度数分布から, 車の数の平均と分散を求めると

$$\bar{x}=\frac{0\times 68+1\times 81+2\times 38+3\times 9+4\times 4}{200}=\mathbf{1}$$
$$s^2=\frac{0^2\times 68+1^2\times 81+2^2\times 83+3^2\times 9+4^2\times 4}{200}-1^2=\mathbf{0.89}$$

平均と分散はほぼ等しいので, 車の交通量の分布は近似的にポアソン分布に従うことが示唆される.

平均 $\bar{x}$ は λ の推定値を与えるから, このデータに当てはめるポアソン分布は,

$$P(X=x)=\frac{e^{-1}}{x!} \qquad (x=0, 1, 2, \cdots)$$

理論度数はこれら確率に 200 をかけて求める.

$$200\times\frac{e^{-1}}{0!}\fallingdotseq 74, \quad 200\times\frac{e^{-1}}{1!}\fallingdotseq 74, \quad 200\times\frac{e^{-1}}{2!}\fallingdotseq 37,$$
$$200\times\frac{e^{-1}}{3!}\fallingdotseq 12, \quad 200\times\frac{e^{-1}}{4!}\fallingdotseq 3$$

よって,

車の数	0	1	2	3	4	計
観測度数	68	81	38	9	4	200
理論度数	74	74	37	12	3	200

［注］　ポアソン分布のデータへの当てはまりが良いか否かは，通常カイ2乗検定によってなされる（9章カイ2乗検定を参照）．

例題 12　**（離散型一様分布）**

乱数表から無作為に選んだ2個の乱数を X，Y とするとき

（a）　$X+Y$　　（b）　$4X-3Y$

の平均と分散を求めよ．

解　X の確率分布は

$$P(X=x)=\frac{1}{10}\qquad (x=0, 1, 2, \cdots, 9)$$

であるから，その平均と分散は

$$E(X)=0\times\frac{1}{10}+1\times\frac{1}{10}+2\times\frac{1}{10}+\cdots+9\times\frac{1}{10}=\frac{1}{10}\times\frac{9(9+1)}{2}=\frac{9}{2}$$

$$V(X)=0^2\times\frac{1}{10}+1^2\times\frac{1}{10}+2^2\times\frac{1}{10}+\cdots+9^2\times\frac{1}{10}-\left(\frac{9}{2}\right)^2$$

$$=\frac{1}{10}\times\frac{9(9+1)(18+1)}{6}-\frac{81}{4}=\frac{33}{4}$$

Y は X と同じ確率分布をもつから

$$E(X)=E(Y)=\frac{9}{2}$$

$$V(X)=V(Y)=\frac{33}{4}$$

よって，

（a）　$E(X+Y)=E(X)+E(Y)=\mathbf{9}$

$V(X+Y)=V(X)+V(Y)$　　（X と Y は独立であるから）

$$=\frac{33}{4}+\frac{33}{4}=\mathbf{\frac{33}{2}}$$

（b）　$E(4X-3Y)=4E(X)-3E(Y)=\mathbf{\frac{9}{2}}$

$$V(4X-3Y)=16V(X)+9V(Y)=25\times\frac{33}{4}=\mathbf{\frac{825}{4}}$$

例題 13　(幾何分布の応用)

ある射撃手の標的への命中率は 0.6 である．この射撃手が標的に命中するまで弾丸を射つとき，射った弾丸の数の平均と標準偏差を求めよ．また，

(a)　射撃手が標的に命中するまでに少なくとも 4 発を射たねばならない確率を求めよ．

(b)　射撃手が標的を命中させるのに n 発を必要とする確率が 0.99 以上になる最小の n を求めよ．

解　標的を最初に命中させるまでの弾丸の数を X とすると，X は $p=0.6$ の幾何分布に従うから

$$P(X=x)=0.6\times(0.4)^{x-1} \qquad (x=1,2,3,\cdots)$$

幾何分布の平均と分散の公式より

$$\mu=\frac{1}{p}=\frac{1}{0.6}\fallingdotseq\mathbf{1.67}$$

$$\sigma=\sqrt{\frac{q}{p^2}}=\sqrt{\frac{1-0.6}{0.6^2}}\fallingdotseq\mathbf{1.05}$$

(a)　$P(X\geqq 4)=1-P(X\leqq 3)$

$$=1-P(1)-P(2)-P(3)$$

$$=1-0.6-0.24-0.096$$

$$=\mathbf{0.064}$$

(b)　$P(X\leqq n)\geqq 0.99$

を満たす最小の n を求めねばならない．

$$P(X\leqq n)=1-P(X\geqq n+1)$$

$$=1-\sum_{x=n+1}^{\infty}0.6(0.4)^{x-1}$$

$$=1-0.4^n$$

だから

$$1-0.4^n\geqq 0.99$$

$$n\geqq\frac{\log 0.01}{\log 0.4}\fallingdotseq 5.03$$

よって，最低 **6 発**が必要．

4章の問題

4.1 (a) サイコロを8回投げるとき，次の確率を求めよ．

(i) 1の目がでない．

(ii) 1の目が4回でる．

(iii) 1の目が少なくとも2回でる．

(b) X が2項分布 $B\left(5, \dfrac{1}{3}\right)$ に従うとき，次の値を求めよ．

(i) $P(X=2)$,

(ii) $P(X \geqq 4)$,

(iii) $P(2 \leqq X \leqq 5)$

4.2 2項分布の平均が16で，分散が3.2のとき，この2項分布の n と p を求めよ．また，この2項分布のモードを求めよ．

4.3 男が生まれる確率は $\dfrac{1}{2}$ であるとして，5人の子供をもつ家族で次の事象の起こる確率を求めよ．

(a) 5人のうち，少なくとも4人が男である．

(b) 男と女が少なくとも1人は含まれる．

(c) 5人とも性別が同じである．

(d) 上2人が女で下3人は男である．

4.4 ある多肢選択型試験は問題が全部で10問あって，各問題は1つの正答を含む4つの選択肢からなる．ある受験生が各問題ごとに答を無作為に選ぶとき，高々2個の正答を得る確率を求めよ．

4.5 サイコロを何回か投げて，6の目が少なくとも1回出る確率を0.95以上にするには，何回の投げが必要か．

4.6 Aは3個の硬貨を投げ，同時にBは4個の硬貨を投げる．そのとき，AがBよりおもてを多く出す確率を求めよ．

4.7 X が2項分布 $B\left(8, \dfrac{1}{2}\right)$ に従うとき，次の値を求めよ．

(a) この分布の平均 μ と標準偏差 σ．

(b) $P(|X-\mu| \geqq 2\sigma)$．

(c) $E\{X(X-3)\}$．

4.8　2項分布 $B(n, p)$ のモードは，$(n+1)p$ が整数でないならば $(n+1)p$ を超えない最大の整数で，$(n+1)p$ が整数ならば $(n+1)p$ と $(n+1)p-1$ の2つであることを示せ．

4.9　600頁のある本には300個の誤字があって，これらは本全体にランダムに分布している．この本の任意の1頁が

（a）　2個の誤字，　（b）　少なくとも2個の誤字

を含む確率を求めよ．

4.10　ある都市における1日当たりの交通事故による死者の数は，平均1.8人のポアソン分布に従うという．このとき次を求めよ．

（a）　この都市のある日の交通事故による死者の数が3人を超える確率．

（b）　この都市のある日の交通事故による死者の数が0人である確率．

4.11　月曜日1時限の講義に遅刻する学生の数は，平均1.2人のポアソン分布に従う．次の確率を求めよ．

（a）　ある週，3人の学生が講義に遅刻する．

（b）　ある週，高々1人の学生が講義に遅刻する．

4.12　確率変数 X がポアソン分布に従い，

$$P(X=3)=5P(X=5)$$

なる関係を満たすとき，次の値を求めよ．

（a）　$P(X=1)$　（b）　$P(X \leqq 3)$

4.13　ある小さなハイヤー会社には5台の車がある．この会社には平日は平均して2台の需要があり，週末には平均して3台の需要がある．車の申込みは1日単位で行われるとして，この会社が次のとき客の申込みを断らねばならなくなる確率を求めよ．

（a）　月曜日　（b）　週末．

4.14　ある機械が作るレンズは平均1.5%が欠陥品である．この機械で作られた100個のレンズの中に

（a）　欠陥品が高々1個含まれる，

（b）　欠陥品が4個以上含まれる

確率を求めよ．

5

正規分布

5-1 正規分布

正規分布　X の確率密度関数が

$$f(x)=\frac{1}{\sqrt{2\pi}\,\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}} \qquad (-\infty<x<\infty)$$

で与えられるとき，この分布を**正規分布**という．式の中の定数 μ と σ^2 は正規分布の平均と分散である．平均 μ，分散 σ^2 の正規分布を $N(\mu,\sigma^2)$ で表し，X が正規分布 $N(\mu,\sigma^2)$ に従うことを

$$X\sim N(\mu,\sigma^2)$$

と書く．

標準正規分布　$\mu=0$，$\sigma=1$ の特別な正規分布 $N(0,1)$ を**標準正規分布**という．その確率密度関数は

$$\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}} \qquad (-\infty<x<\infty)$$

で与えられる．

標準化　$X\sim N(\mu,\sigma^2)$ のとき，

$$Z=\frac{X-\mu}{\sigma}$$

は標準正規分布 $N(0,1)$ に従う．X から Z へのこの変換を**標準化**といい，Z を**標準化変量**という．

正規分布の平均と分散

$$E(X)=\mu$$
$$V(X)=\sigma^2$$

正規分布表　Z が $N(0,1)$ に従うとき，Z の累積分布関数

$$\Phi(z)=P(Z<z)=\int_{-\infty}^{z}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx$$

の値を $z(>0)$ の各値に対して表にしたものを**正規分布表**という（図 1 参照）．$-z$ に対する $\Phi(z)$ の値は，

$$\Phi(-z)=1-\Phi(z)$$

から求まる（図 2 参照）．

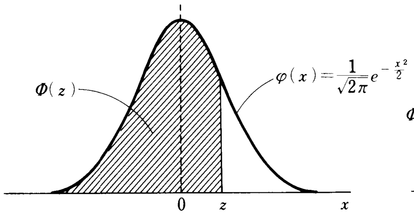


図 1

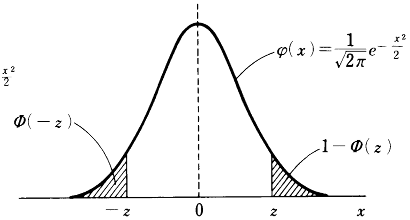


図 2

2 項分布の正規近似　X が 2 項分布 $B(n,p)$ に従うとき，n が十分大きいならば，X の分布は正規分布 $N(np, npq)$ で近似される．したがって，n が大きいならば

$$Z=\frac{X-np}{\sqrt{npq}}$$

の分布は $N(0,1)$ で近似される．

実際には，n と p が

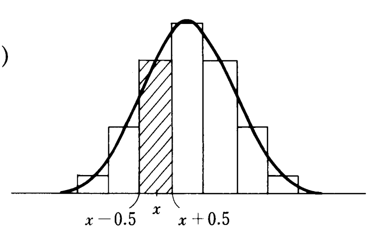
$$np>5 \quad かつ \quad nq>5$$

を満たすとき，2 項分布の正規分布への近似は十分とされる．

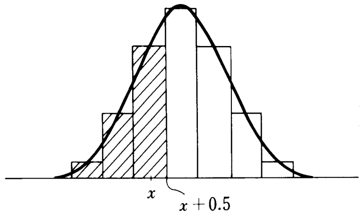
5-2 半整数補正

X は整数値のみをとる離散型確率変数で，Y は連続型確率変数とする．2 項分布を正規分布で近似するときのように，離散変量 X の分布を連続変量 Y の分布で近似して確率の計算をするとき，

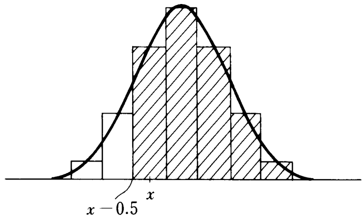
$$P(X=x)\cong P(x-0.5<Y<x+0.5)$$



$$P(X \leqq x) \cong P(Y < x + 0.5)$$

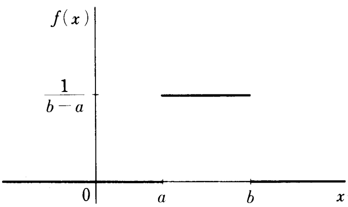


$$P(X \geqq x) \cong P(Y > x - 0.5)$$



のように近似することを**半整数補正**（または**連続性の補正**）という．

5-3 その他の連続型分布



連続型一様分布　X の確率密度関数が

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

で与えられるとき，この分布を**連続型一様分布**（または**矩形分布**）という．

連続型一様分布の平均と分散

$$\mu=\frac{a+b}{2}$$

$$\sigma^2=\frac{(b-a)^2}{12}$$

指数分布　X の確率密度関数が

$$f(x)=\begin{cases} \theta e^{-\theta x} & (x \geqq 0,\ \theta>0) \\ 0 & (\text{その他}) \end{cases}$$

で与えられるとき，この分布を母数 θ の**指数分布**という．

指数分布の平均と分散

$$\mu=\frac{1}{\theta}$$

$$\sigma^2=\frac{1}{\theta^2}$$

例　題

例題 1　(正規分布表の使い方)

Z が $N(0,1)$ に従うとき，正規分布表によって次の値を求めよ．

(a)　$P(Z<1)$

(b)　$P(Z<1.24)$

(c)　$P(Z<-0.5)$

(d)　$P(-1<Z<0.5)$

(e)　$P(1.51<Z<2.16)$

(f)　$P(-1.64<Z<-0.8)$

(g)　$P(Z<c)=0.65$ を満たす c の値

(h)　$P(Z>c)=0.42$ を満たす c の値

解　正規分布表を正確に使うためには，必要に応じて図をかいてみるのがよい．

以下において $\Phi(z)$ は Z の累積分布関数を表す．すなわち，

$$\Phi(z)=P(Z\leqq z)=P(Z<z)$$

正規変数の確率計算では，不等式に等号があってもなくても同じ値となるので，本書では等号をつけていない．

(a)　$P(Z<1)=\Phi(1)=\mathbf{0.8413}$

(b)　$P(Z<1.24)=\Phi(1.24)=\mathbf{0.8925}$

(c)　$P(Z<-0.5)=\Phi(-0.5)=1-\Phi(0.5)=1-0.6915=\mathbf{0.3085}$

(d)　$P(-1<Z<0.5)=P(Z<0.5)-P(Z<-1)$

$$=\Phi(0.5)-\{1-\Phi(1)\}$$

$$=0.6915+0.8413-1=\mathbf{0.5328}$$

(e)　$P(1.51<Z<2.16)=\Phi(2.16)-\Phi(1.51)$

$$=0.9846-0.9345=\mathbf{0.0501}$$

(f)　$P(-1.64<Z<-0.8)=P(0.8<Z<1.64)$　(分布の対称性より)

$$=\Phi(1.64)-\Phi(0.8)$$

$$=0.9495-0.7881=\mathbf{0.1614}$$

(g)　$P(Z<c)=0.65$ を満たす c は，正規分布表より

$$\Phi(0.38)=0.6480,\qquad \Phi(0.39)=0.6517$$

であるから，補間によって

$$c=0.38+\frac{0.65-0.6480}{0.6517-0.6480}(0.39-0.38)$$
$$=0.38+0.005=\mathbf{0.385}$$

(h)　$P(Z>c)=0.42$

$P(Z<c)=0.58$

表より

$$\Phi(0.20)=0.5793,\quad \Phi(0.21)=0.5832$$

であるから，補間によって，

$$c=0.20+\frac{0.58-0.5793}{0.5832-0.5793}(0.21-0.20)$$
$$=\mathbf{0.202}$$

例題 2　**（正規分布表の使い方）**

X が $N(10, 2^2)$ に従うとき，次の値を求めよ．

(a)　$P(X<13)$

(b)　$P(X>11)$

(c)　$P(X>8)$

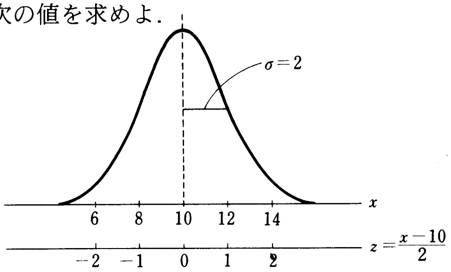
(d)　$P(X<7)$

(e)　$P(9<X<12)$

(f)　$P(7.8<X<9.6)$

(g)　$P(X<c)=0.85$ を満たす c の値

(h)　$P(X<c)=0.3$ を満たす c の値



解　$X \sim N(10, 2^2)$ より

$$Z=\frac{X-10}{2}\sim N(0,1)$$

よって，

(a)　$P(X<13)=P\left(Z<\frac{13-10}{2}\right)=\Phi(1.5)=\mathbf{0.9332}$

(b)　$P(X>11)=1-P(X<11)$

$$=1-P\left(Z<\frac{11-10}{2}\right)$$
$$=1-\Phi(0.5)=1-0.6915=\mathbf{0.3085}$$

(c)　$P(X>8)=P\left(Z>\frac{8-10}{2}\right)$

$$=P(Z>-1)$$

$$=1-P(Z<-1)=1-\phi(-1)=\phi(1)=\mathbf{0.8413}$$

(d)　$P(X<7)=P\left(Z<\dfrac{7-10}{2}\right)$

$$=P(Z<-1.5)$$

$$=\Phi(-1.5)$$

$$=1-\Phi(1.5)=1-0.9332=\mathbf{0.0668}$$

(e)　$P(9<X<12)=P\left(\dfrac{9-10}{2}<Z<\dfrac{12-10}{2}\right)$

$$=P(-0.5<Z<1)$$

$$=P(Z<1)-P(Z<-0.5)$$

$$=\Phi(1)-\Phi(-0.5)$$

$$=\Phi(1)+\Phi(0.5)-1$$

$$=0.8413+0.6915-1=\mathbf{0.5328}$$

(f)　$P(7.8<X<9.6)=P\left(\dfrac{7.8-10}{2}<Z<\dfrac{9.6-10}{2}\right)$

$$=P(-1.1<Z<-0.2)$$

$=P(0.2<Z<1.1)$　（分布の対称性より）

$$=P(Z<1.1)-P(Z<0.2)$$

$$=\Phi(1.1)-\Phi(0.2)$$

$$=0.8643-0.5793=\mathbf{0.2850}$$

(g)　$P(X<c)=P\left(Z<\dfrac{c-10}{2}\right)=\Phi\left(\dfrac{c-10}{2}\right)$

より

$$\Phi\left(\frac{c-10}{2}\right)=0.85$$

を満たす c を求めればよい．正規分布表より

$$\Phi(1.03)=0.8485,\qquad \Phi(1.04)=0.8508$$

補間によって，

$$\frac{c-10}{2}=1.03+\frac{0.85-0.8485}{0.8508-0.8485}\times 0.01\fallingdotseq 1.037$$

$$c=10+2\times 1.037\fallingdotseq \mathbf{12.07}$$

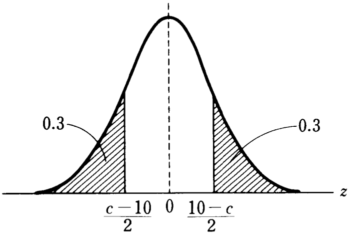
(h)　$P(X<c)=\Phi\left(\dfrac{c-10}{2}\right)$ より，c は

$$\Phi\left(\frac{c-10}{2}\right)=0.3$$

の解である. しかし, 右辺の 0.3 は 0.5 より小さいから, 右図より $\dfrac{c-10}{2}$ の値は負になる. よって, $\Phi(z)$ $(z>0)$ の表を使うには, 図からわかるように, 上の式に代って

$$\Phi\left(\frac{10-c}{2}\right)=0.7$$

から c を求めねばならない.



表より, $\Phi(0.52)=0.6985$, $\Phi(0.53)=0.7019$ であるから補間によって

$$\frac{10-c}{2}=0.52+\frac{0.7-0.6985}{0.7019-0.6985}\times 0.01 \fallingdotseq 0.5234$$

$$c=10-2\times 0.5234 \fallingdotseq \mathbf{8.953}$$

例題 3　(正規分布の応用)

測定器具である物の長さを測るときの誤差は, 平均 0, 標準偏差 0.2 mm の正規分布に従う. この器具による 1 回の測定の誤差が

(a)　0.5 mm 以上,

(b)　0.3 mm 以内

となる確率を求めよ.

解　測定値の誤差を X とすると

$$X \sim N(0, 0.2^2)$$

であるから,

(a)　$P(|X|>0.5)$

$=2P(X>0.5)$　(正規分布の対称性より)

$=2P\left(Z>\dfrac{0.5-0}{0.2}\right)$

$=2P(Z>2.5)$

$=2(1-P(Z<2.5))$

$=2(1-\Phi(2.5))=2(1-0.9938)=\mathbf{0.0124}$

(b)　$P(|X|<0.3)$

$=P(-0.3<X<0.3)$

$=P\left(\dfrac{-0.3-0}{0.2}<Z<\dfrac{0.3-0}{0.2}\right)$

$=P(-1.5<Z<1.5)$

$$=\Phi(1.5)-\Phi(-1.5)$$
$$=2\Phi(1.5)-1=2\times 0.9332-1=\mathbf{0.8664}$$

例題 4 (正規分布の応用)

IQ が $N(100, 225)$ に従うとき，IQ が

(a) 85 以下，　(b) 90 から 120 の間，　(c) 130 以上

の人の割合を求めよ．また，上位 10% にある IQ の最小値を求めよ．

解　IQ を X とすると

$$X\sim N(100, 15^2)$$

よって

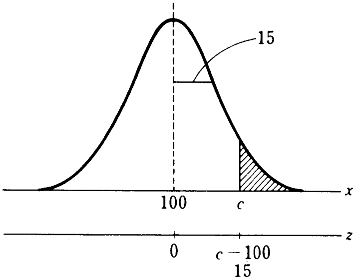
(a)　$P(X<85)=P\left(Z<\dfrac{85-100}{15}\right)$
$$=\Phi(-1)=1-\Phi(1)=1-0.8413=\mathbf{0.1587}$$

(b)　$P(90<X<120)=P\left(\dfrac{90-100}{15}<Z<\dfrac{120-100}{15}\right)$
$$=\Phi(1.33)-\Phi(-0.67)$$
$$=\Phi(1.33)+\Phi(0.67)-1$$
$$=0.9082+0.7486-1=\mathbf{0.6568}$$

(c)　$P(X>130)=1-P(X<130)$
$$=1-P\left(Z<\dfrac{130-100}{15}\right)$$
$$=1-\Phi(2)=1-0.9772=\mathbf{0.0228}$$

上位 10% の IQ の最小値は右の図の c の値を求めればよいから

$$0.10=P(X>c)$$
$$=P\left(Z>\dfrac{c-100}{15}\right)$$
$$=1-\Phi\left(\dfrac{c-100}{15}\right)$$



よって，

$$\Phi\left(\dfrac{c-100}{15}\right)=0.90$$

正規分布表より

$$\Phi(1.28)=0.8997,\qquad \Phi(1.29)=0.9015$$

であるから，補間によって

$$\frac{c-100}{15}=1.28+\frac{0.90-0.8997}{0.9015-0.8997}\times 0.01 \fallingdotseq 1.282$$

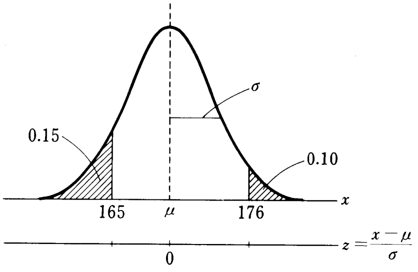
$$c=100+15\times 1.282=\mathbf{119.2}$$

例題 5 （正規分布の応用）

高校3年生の男子の身長の分布は正規分布に従うことが知られている．これら生徒の 10%はその身長が 176 cm を超え，15%は 165 cm 以下である．男子高校3年生の身長の平均と標準偏差を求めよ．

解　高校3年生の男子の身長を X とし，その平均と標準偏差を μ と σ で表すと，

$$X \sim N(\mu, \sigma^2)$$



与えられた条件より

$$P(X>176)=0.10$$

$$P\left(Z>\frac{176-\mu}{\sigma}\right)=0.10$$

$$P\left(Z<\frac{176-\mu}{\sigma}\right)=0.90$$

正規分布表より

$$\frac{176-\mu}{\sigma}=1.282 \qquad \cdots(1)$$

もう1つの条件から

$$P(X<165)=0.15$$

$$P\left(Z<\frac{165-\mu}{\sigma}\right)=0.15$$

正規分布表より

$$\frac{165-\mu}{\sigma}=-1.037 \quad \cdots(2)$$

(1)，(2) より

$$\mu+1.282\sigma=176$$

$$\mu-1.037\sigma=165$$

これを解いて

$$\mu=\mathbf{169.9}\ \text{(cm)}$$

$$\sigma=\ \mathbf{4.7}\ \text{(cm)}$$

例題 6 （正規分布の応用）

300 人の学生の「統計学」の試験の結果から，その得点分布は近似的に平均 55 点，標準偏差 10 点の正規分布に従うとみなされた．

（a） この試験で得点が 60 点から 70 点までの人数は約何人いるか．

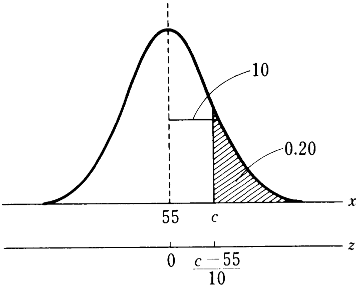
（b） 成績が上位のもの 20%に‘優’をつけるとき，何点以上が優になるか．

解 得点を X とすると，$X \sim N(55, 10^2)$．
よって

（a） $P(60<X<70)=P\left(\dfrac{60-55}{10}<Z<\dfrac{70-55}{10}\right)$

$$=\Phi(1.5)-\Phi(0.5)=0.9332+0.6915=0.2417$$

$300\times 0.2417=72.51$．よって約 **73** 人．



（b） 求める点数を c とすると

$P(X>c)=0.2$

$P(X<c)=0.8$

$P\left(Z<\dfrac{c-55}{10}\right)=0.8$

正規分布表より

$\Phi(0.84)=0.7995$,　　$\Phi(0.85)=0.8023$

よって，補間により

$$\frac{c-55}{10}=0.84+\frac{0.8-0.7995}{0.8023-0.7995}\times 0.01\fallingdotseq 0.842$$

$$c=55+10\times 0.842=63.42$$

ゆえに，**64 点以上**が優である．

例題 7 （正規分布の応用）

ある走り幅跳び選手の飛距離 X は平均 6.5 m，標準偏差 0.2 m の正規分布に従う．（a） この選手の 1 回の飛距離が 6.9 m を超える確率を求めよ．（b） この選手が 3 回跳ぶとき，3 回中 1 回だけ飛距離が 6.9 m を超える確率を求めよ．（c） 100 回に 1 回，この選手が超えると期待される飛距離はいくらか．

解 $X \sim N(6.5, 0.2^2)$

（a）　$P(X>69)=1-P(X<6.9)$

$$=1-P\left(Z<\frac{6.9-6.5}{0.2}\right)$$

$$=1-\Phi(2)=1-0.9772=\mathbf{0.0228}$$

（b）　P(3 回中 1 回，X が 6.9 を超える)$={}_3C_1(0.0228)(0.9772)^2=\mathbf{0.0653}$

（c）　期待される値を c とすると

$$P(X>c)=0.01$$

$$P(X<c)=0.99$$

$$P\left(Z<\frac{c-6.5}{0.2}\right)=0.99$$

$$\Phi\left(\frac{c-6.5}{0.2}\right)=0.99$$

正規分布表より，

$$\Phi(2.32)=0.9898, \qquad \Phi(2.33)=0.9901$$

よって，補間により

$$\frac{c-6.5}{0.2}=2.32+\frac{0.99-0.9898}{0.9901-0.9898}\times 0.01 \fallingdotseq 2.327$$

$$c=6.5+0.2\times 2.327 \fallingdotseq \mathbf{6.97}\ (\mathrm{m})$$

例題 8　（正規分布の応用）

時刻表によると，毎日ある駅に午前 9 時 30 分に刻着する列車がある．延べ 10 日間にわたって，この列車の定刻からの遅れ（分）を調べ，次の結果を得た．

3, 0, 10, −1, 6, 8, −2, 5, 0, 1

この列車の到着時刻の平均と標準偏差を求めよ．列車の到着時刻はこれと同じ平均，同じ標準偏差の正規分布に従うと仮定して，ある日，列車が

（a）　定刻より 10 分以上遅れる確率，

（b）　定刻より早く到着する確率

を求めよ．

解　到着時刻の平均を $\bar{x}$，標準偏差を s とすると

$$\bar{x}=\frac{3+0+10-1+6+8-2+5+0+1}{10}=\mathbf{3}\ (分)$$

$$s=\sqrt{\frac{3^2+0^2+10^2+(-1)^2+6^2+8^2+(-2)^2+5^2+0^2+1^2}{10}-3^2}=\sqrt{15}\fallingdotseq\mathbf{3.87}\ (分)$$

この列車の駅への到着時刻を X とすると，仮定より

$$X \sim N(3, 15)$$

よって

（a） P(定刻より 10 分以上遅れる)

$$= P(X > 10)$$

$$= P\left(Z > \frac{10-3}{\sqrt{15}}\right)$$

$$= P(Z > 1.81)$$

$$= 1 - \Phi(1.81) = 1 - 0.9649 = \mathbf{0.0351}$$

（b） P(定刻より早く到着する)

$$= P(X < 0)$$

$$= P\left(X < \frac{0-3}{\sqrt{15}}\right)$$

$$= P(X < -0.77)$$

$$= \Phi(-0.77) = 1 - \Phi(0.77) = 1 - 0.7794 = \mathbf{0.2206}$$

例題 9 （2 項分布の正規近似）

硬貨を 400 回投げるとき，おもてが 180 回から 210 回まで出る確率を求めよ．

解 X をおもての出る数とすると，$X \sim B\left(400, \frac{1}{2}\right)$

$$n = 400, \quad np = nq = 400 \times \frac{1}{2} = 200 \geqq 5, \quad npq = 400 \times \frac{1}{2} \times \frac{1}{2} = 100$$

$n = 400$ は十分大きいから，2 項分布 $B\left(400, \frac{1}{2}\right)$ は正規分布 $N(200, 10^2)$ で近似できる．

Y を平均が 200 で，分散が 10^2 の正規変量とすると，

$$Z = \frac{Y - 200}{10} \sim N(0, 1)$$

よって求める確率は，半整数補正により

$$P(180 \leqq X \leqq 210) \fallingdotseq P(179.5 < Y < 210.5)$$

$$= P\left(\frac{179.5 - 200}{10} < Z < \frac{210.5 - 200}{10}\right)$$

$$= P(-2.05 < Z \leq 1.05)$$

$$= \Phi(1.05) + \Phi(2.05) - 1$$
$$= 0.8531 + 0.9798 - 1 = \mathbf{0.8329}$$

例題 10　(2 項分布の正規近似)

ある人があるゲームに勝つ確率を $\frac{1}{3}$, 負ける確率を $\frac{2}{3}$ とする．ゲームに勝てば 1000 円得をし，負ければ 250 円損をする．この人がこのゲームを 20 回行うとき，少なくとも 3000 円の得をする確率を求めよ．

解　この人がゲームに勝つ回数を x とすると，20 回のゲームによるこの人の利益は

$$1000x - 250(20 - x)$$

で，これが 3000 より大きいことから

$$1000x - 250(20 - x) \geqq 3000$$
$$x \geqq 6.4$$

よって，少なくとも 3000 円の得をするには，20 回中 7 回以上ゲームに勝たねばならない．その確率は $n = 20$, $p = \frac{1}{3}$ の 2 項分布より

$$\sum_{x=7}^{20} {}_{20}\mathrm{C}_x \left(\frac{1}{3}\right)^x \left(\frac{2}{3}\right)^{20-x}$$

この確率の計算には 2 項分布の正規近似を使う．

$$\mu = np = 20 \times \frac{1}{3} \fallingdotseq 6.67,$$

$$\sigma = \sqrt{npq} = \sqrt{20 \times \frac{1}{3} \times \frac{2}{3}} \fallingdotseq 2.11,$$

$$Z = \frac{X - \mu}{\sigma} = \frac{X - 6.67}{2.11},$$

$$\frac{6.5 - 6.67}{2.11} = -0.08, \qquad \frac{20.5 - 6.67}{2.11} = 6.55$$

であるから，

$$P(X \geqq 7) = \sum_{x=7}^{20} {}_{20}\mathrm{C}_x \left(\frac{1}{3}\right)^x \left(\frac{2}{3}\right)^{20-x}$$
$$\fallingdotseq \int_{-0.08}^{6.55} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz \fallingdotseq \Phi(0.08) = \mathbf{0.532}$$

例題 11　**(正規分布とポアソン分布)**

郵便配達員が月曜日の朝，ある家に配達に行く時刻 T は，平均午前9時50分，標準偏差10分の正規分布に従い，配達する郵便物の数 X は平均3のポアソン分布に従う．このとき，次の確率を求めよ．

(a)　この家が月曜日の朝1通の郵便物を受け取る．

(b)　この家の主人が午前10時に家を出た後に配達員がくる．

(c)　配達員が午前8時50分から午前9時55分の間に，その家に3通以上の郵便物を届ける．

解　与えられた情報から，

$$T \sim N(9:50, 10^2)$$

$$X \sim P(X=x)=\frac{3^x e^{-3}}{x!} \qquad (x=0, 1, 2, \cdots)$$

よって，

(a)　$P(X=1)=3e^{-3}=\mathbf{0.15}$

(b)

$$\begin{aligned} P(T>10:00) &= P\left(Z>\frac{10:00-9:50}{10}\right) \\ &= P(Z>1) \\ &= 1-\Phi(1) \\ &= 1-0.8413=\mathbf{0.1587} \end{aligned}$$

(c)　$P\{(8:50<T<9:55)\cap(X\geqq 3)\}$

$$=P(8:50<T<9:55)P(X\geqq 3) \qquad (T \text{ と } X \text{ は独立だから})$$

$$=P\left(\frac{8:50-9:50}{10}<Z<\frac{9:55-9:50}{10}\right)(1-P(0)-P(1)-P(2))$$

$$=P\left(-6<Z<\frac{1}{2}\right)(1-e^{-3}-3e^{-3}-4.5e^{-3})$$

$$=0.6915(1-0.423)=\mathbf{0.399}$$

例題 12　**(指数分布の平均と分散)**

指数分布

$$f(x)=\begin{cases} \theta e^{-\theta x} & (x\geqq 0,\ \theta>0) \\ 0 & (\text{その他}) \end{cases}$$

の平均と分散を求めよ．

解　$E(X)=\theta\int_0^\infty xe^{-\theta x}\,dx$

$$=\theta\left[\frac{-xe^{-\theta x}}{\theta}\right]_0^\infty+\int_0^\infty e^{-\theta x}\,dx \qquad \text{(部分積分による)}$$

$$=0+\int_0^\infty e^{-\theta x}\,dx=\left[-\frac{e^{-\theta x}}{\theta}\right]_0^\infty=\boldsymbol{\frac{1}{\theta}}$$

$$E(X^2)=\theta\int_0^\infty x^2e^{-\theta x}\,dx$$

$$=\theta\left[\frac{-x^2e^{-\theta x}}{\theta}\right]_0^\infty+2\int_0^\infty xe^{-\theta x}\,dx \qquad \text{(部分積分による)}$$

$$=0+\frac{2}{\theta}E(X)=\frac{2}{\theta^2}$$

よって,

$$V(X)=E(X^2)-\{E(X)\}^2$$

$$=\frac{2}{\theta^2}-\frac{1}{\theta^2}=\boldsymbol{\frac{1}{\theta^2}}$$

例題 13　(連続型一様分布)

長さ a の線分上でランダムに1点を選ぶ. 短かい方の線分と長い方の線分の長さの比が $\frac{1}{3}$ より小さい確率を求めよ.

解　与えられた線分の左端から選ばれた点までの距離を X とすると, 残りの部分の長さは $a-X$ である. 仮定より, X は一様分布

$$f(x)=\begin{cases}\dfrac{1}{a} & (0<x<a)\\[2mm] 0 & (\text{その他})\end{cases}$$

に従う. ところで

$$X<\frac{a}{2}\text{ のとき, }\frac{X}{a-X}<\frac{1}{3} \quad\Rightarrow\quad X<\frac{a}{4}$$

$$X>\frac{a}{2}\text{ のとき, }\frac{a-X}{X}<\frac{1}{3} \quad\Rightarrow\quad X>\frac{3}{4}a$$

となるから, 短い方の線分と長い方の線分の長さの比を R とすると

$$P\left(R<\frac{1}{3}\right)=P\left(X<\frac{a}{2}\right)P\left(X<\frac{a}{4}\,\middle|\,X<\frac{a}{2}\right)+P\left(X>\frac{a}{2}\right)P\left(X>\frac{3}{4}a\,\middle|\,X>\frac{a}{2}\right)$$

ここで

$$P\left(X<\frac{a}{4}\,\middle|\,X<\frac{a}{2}\right)=\frac{P\left\{\left(X<\frac{a}{4}\right)\cap\left(X<\frac{a}{2}\right)\right\}}{P\left(X<\frac{a}{2}\right)}=\frac{P\left(X<\frac{a}{4}\right)}{P\left(X<\frac{a}{2}\right)}$$

同様にして,

$$P\left(X>\frac{3}{4}a\,\middle|\,X>\frac{a}{2}\right)=\frac{P\left(X>\frac{3}{4}a\right)}{P\left(X>\frac{a}{2}\right)}$$

よって,

$$P\left(R<\frac{1}{3}\right)=P\left(X<\frac{a}{4}\right)+P\left(X>\frac{3}{4}a\right)$$
$$=\frac{a}{4}\cdot\frac{1}{a}+\frac{a}{4}\cdot\frac{1}{a}=\mathbf{\frac{1}{2}}$$

5章の問題

5.1 Z が $N(0,1)$ に従うとき，正規分布表によって次の値を求めよ.

(a) $P(Z<1.64)$.

(b) $P(Z>1.15)$.

(c) $P(Z<-0.34)$.

(d) $P(-1<Z<0.5)$.

(e) $P(1.24<Z<2.16)$.

(f) $P(Z>-2.19)$.

(g) $P(Z>c)=0.38$ を満たす c の値.

(h) $P(Z<c)=0.19$ を満たす c の値.

5.2 X が $N(3,1)$ に従うとき，次の値を求めよ.

(a) $P(X<3)$.

(b) $P(X<4.93)$.

(c) $P(X>3.06)$.

(d) $P(1<X<4.2)$.

(e) $P(1.5<X<2.5)$.

(f) $P(|X-2|<1)$.

(g) $P(X>c)=0.1$ を満たす c の値.

(h) $P(X<c)=0.2$ を満たす c の値.

5.3 ある種の電球の寿命 X は，過去の経験から平均1500時間，標準偏差25時間の正規分布に従うことが知られている．寿命 X が

（a） 1530時間以上，

（b） 1480時間未満，

（c） 1475時間から1550時間の間

にある電球の割合を求めよ．

5.4 軍隊で使われる靴下の寿命は平均55日，標準偏差8日の正規分布に従うといわれている．ある日，5000人の兵士に靴下を与えたとき，45日以内には何足を補給しなければならないか．また，61日以内ではどうか．

5.5 ある機械が作る部品の長さは標準偏差が2 cmの正規分布に従う．（a）これら部品の97.5%はその長さが7.5 cm以下であるとき，部品の長さの平均を求めよ．また，（b）この機械が作る部品の長さが5.4 cmから5.5 cmの間にある確率を求めよ．

5.6 自動充塡機によってある食品を正味50 g入りと書かれた袋に詰める．機械が1袋に詰める実際の重さは平均52.5 g，標準偏差1.6 gの正規分布に従うことがわかっているとき，（a）袋の中の食品の重さが50 gを下回る確率はいくらか．（b）この確率を1%以下にするには，機械が詰める食品の重さの平均をいくらに定めればよいか．

5.7 以下の数値はある生徒の10日間の通学時間（分）を示したものである．

36　32　26　22　44　38　34　32　42　34

通学時間の平均と標準偏差を求めよ．

この生徒の通学時間はこれら平均と標準偏差をもつ正規分布に従うとして，（a）生徒のある日の通学時間が38分以上となる確率を求めよ．（b）通学時間がある時間を超えることは高々10回に1回にしたいとすれば，その時間は何分か．

5.8 正規分布の密度関数

$$f(x)=\frac{1}{\sqrt{2\pi}\,\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}} \qquad (-\infty<x<\infty)$$

の変曲点の x 座標は $\mu\pm\sigma$ であることを示せ．

5.9 硬貨を 12 回投げるとき，おもてが 9 回以上出る確率を

（a） 2 項分布を用いて，

（b） 2 項分布の正規近似を用いて

求めよ．

5.10 X が $B(50, 0.4)$ に従うとき，次の確率を求めよ．

（a） $X=20$

（b） $15 \leqq X \leqq 25$

5.11 過去の経験から，あるレストランではテーブル予約客の 8 人に 1 人は当日現れないという．このレストランの収容定員は 45 人であるが，毎晩 50 人までの予約を受けつけている．このレストランが当日現れた客のすべてを収容できる確率を求めよ．

5.12 X が $N(0, \sigma^2)$ に従うとき，$|X|$ の平均と分散を求めよ．

5.13 X が区間 $-1 \leqq x \leqq 1$ で一様分布をするとき，次を求めよ．

（a） $P\left(X > \frac{2}{3} \middle| |X| > \frac{1}{2}\right)$

（b） $P\left(X^2 \leqq \frac{1}{4}\right)$

6

無作為抽出と標本分布

6-1　無作為抽出

母集団と標本　調査や実験で観測の対象となる同種の事物の集まりを母集団という．母集団にはそれを構成する要素の数が有限か無限かによって，**有限母集団**と**無限母集団**がある．母集団に関する情報を得るため，それからとり出された母集団の一部分を**標本**という．

無作為抽出　標本から母集団に関する統計的推測を行うためには，標本は母集団の縮図になるようなものでなければならない．そのような標本は無作為抽出によって得られる．**無作為抽出**とは，母集団を構成するすべての要素が等確率で標本のなかに選ばれるような標本抽出の方法である．無作為抽出によって得られた標本を**無作為標本**(または単に**標本**)という．実際の無作為抽出では**乱数表**がよく使われる．

復元抽出と非復元抽出　母集団から標本を抽出するとき，一度とり出したものを元に戻し，次のものをとり出す方法を**復元抽出**といい，とり出したものを元に戻さず，次のものをとり出す方法を**非復元抽出**という．

母数と統計量　母集団におけるある変量 X の確率分布をその変量の**母集団分布**といい，この分布の特性値である平均 μ，分散 σ^2 などを**母数**という．一般に，母集団からとられた大きさ n の無作為標本を表す確率変数 $X_1, X_2, \cdots, X_n$ を**標本変量**という．標本変量 $X_1, X_2, \cdots, X_n$ は互いに独立な n 個の確率変数で，その確率分布はすべて母集団分布と同じである．標本の個数 n を**標本の大きさ**といい，

標本平均　$\overline{X}=\dfrac{1}{n}\sum_{i=1}^{n} X_i$

標本分散　$s^2=\dfrac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2$

のような標本変量 $X_1, X_2, \cdots, X_n$ の関数を**統計量**という．

6-2　標本平均の分布

標本平均 $\overline{X}$ の分布に関して，以下の定理が成り立つ．

$\overline{X}$ の平均と分散

定理 1　平均 μ，分散 σ^2 の無限母集団からとられた大きさ n の標本の平均を $\overline{X}$ とすれば

$$E(\overline{X})=\mu$$

$$V(\overline{X})=\frac{\sigma^2}{n}$$

定理 2　平均 μ，分散 σ^2 の有限母集団からとられた大きさ n の標本の平均を $\overline{X}$，母集団の大きさを N とすれば

$$E(\overline{X})=\mu$$

$$V(\overline{X})=\frac{N-n}{N-1}\cdot\frac{\sigma^2}{n}$$

$\overline{X}$ の標本分布

定理 3　平均 μ，分散 σ^2 の正規母集団からとられた大きさ n の標本の平均 $\overline{X}$ の分布は，平均 μ，分散 $\dfrac{\sigma^2}{n}$ の正規分布に従う．

定理 4（中心極限定理）　平均 μ，分散 σ^2 のある母集団からとられた大きさ n の標本の平均 $\overline{X}$ の分布は，n が十分大きいならば，近似的に平均 μ，分散 $\dfrac{\sigma^2}{n}$ の正規分布に従う．

6-3　χ^2 分布，t 分布，F 分布

χ^2 分布，t 分布，F 分布はいずれも正規母集団からの標本抽出に関連して導出された標本分布である．

χ^2 分布　確率密度関数が

$$f_\nu(x)=ce^{-\frac{x}{2}}x^{\frac{\nu}{2}-1}\qquad(x>0)$$

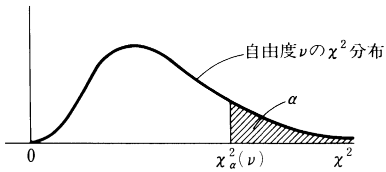
で与えられる分布を**自由度 ν の χ^2 分布**という．ここで，定数 c は $f_\nu(x)$ が密度関数であるという条件から決まる．

χ^2 分布表　確率変数 X が自由度 ν の χ^2 分布をするとき，確率 α に対し

て

$$P(X > \chi_\alpha^2(\nu)) = \alpha$$

を満たす $\chi_\alpha^2(\nu)$ の値の表を χ^2 分布表という（付表 5 参照）．



定理 5　平均 μ，分散 σ^2 の正規母集団からとられた大きさ n の標本の分散を s^2 とするとき，

$$\chi^2 = \frac{ns^2}{\sigma^2}$$

は自由度 $\nu = n-1$ の χ^2 分布に従う．

***t* 分布**　確率密度関数が

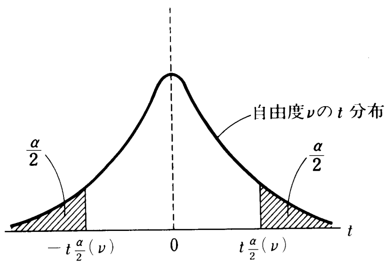
$$g_\nu(x) = c\left(1+\frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}} \qquad (-\infty < x < \infty)$$

で与えられる分布を自由度 ν の ***t* 分布**という．ここで，定数 c は $g_\nu(t)$ が密度関数であるという条件から決まる．

***t* 分布表**　確率変数 T が自由度 ν の t 分布をするとき，確率 α に対して

$$P(|T| > t_{\frac{\alpha}{2}}(\nu)) = \alpha$$

を満たす $t_{\frac{\alpha}{2}}(\nu)$ の値の表を t 分布表という（付表 4 参照）．



定理 6　正規母集団からとられた大きさ n の標本の平均と分散を，$\bar{x}$，s^2 とするとき，

$$t = \frac{\sqrt{n-1}(\bar{x}-\mu)}{s}$$

の分布は，自由度 $\nu = n-1$ の t 分布に従う．

定理 7　2 つの正規母集団 $N(\mu_1, \sigma^2)$，$N(\mu_2, \sigma^2)$ から独立にとられた大きさ n_1，n_2 の標本の平均と分散を $\bar{x}_1$，$\bar{x}_2$ および $s_1{}^2$，$s_2{}^2$ とする．このとき，

$$t = \frac{\bar{x}_1 - \bar{x}_2 - (\mu_1 - \mu_2)}{s\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}$$

ここで，

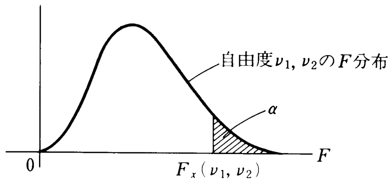
$$s^2 = \frac{n_1 s_1{}^2 + n_2 s_2{}^2}{n_1 + n_2 - 2}$$

は，自由度 $\nu = n_1 + n_2 - 2$ の t 分布に従う．

F 分布　確率密度関数が

$$h_{\nu_1,\nu_2}(x)=cx^{\frac{\nu_1}{2}-1}\left(1+\frac{\nu_1}{\nu_2}x\right)^{-\frac{\nu_1+\nu_2}{2}} \qquad (x>0)$$

で与えられる分布を自由度 ν_1, ν_2 の F 分布という．ここで定数 c は $h_{\nu_1,\nu_2}(x)$ が密度関数であるという条件から決まる．



F 分布表　確率変数 F が自由度 $(\nu_1,\ \nu_2)$ の F 分布をするとき，確率 α に対して，

$$P(F>F_\alpha(\nu_1,\ \nu_2))=\alpha$$

を満たす $F_\alpha(\nu_1,\ \nu_2)$ の値の表を F 分布表という．本書では，$\alpha=0.05$, 0.025 に対する表を与える(付表 6, 付表 7 参照)．

定理 8　2 つの正規母集団 $N(\mu_1,\ \sigma^2)$, $N(\mu_2,\ \sigma^2)$ から独立にとられた n_1 個と n_2 個の標本の不偏分散を $u_1{}^2$, $u_2{}^2$ とするとき，

$$F=\frac{u_1{}^2}{u_2{}^2}$$

の分布は，自由度 $(n_1-1,\ n_2-1)$ の F 分布に従う．

例　　題

例題 1　(非復元抽出による $\overline{X}$ の標本分布)

5 個の数，$\{1, 2, 3, 4, 5\}$ からなる母集団がある．(a) この母集団の平均と分散を求めよ．(b) この母集団から大きさ 2 の標本を非復元抽出でとり出すとき，すべての可能な標本を列挙せよ．(c) 各標本の平均を求め，標本平均 $\overline{X}$ の標本分布を導け．(d) この分布の平均と分散を求めて，これらの値が公式から求めた値と一致することを示せ．(e) 母集団分布と $\overline{X}$ の標本分布をそれぞれ図示せよ．

解　(a)　母平均：$\mu=\dfrac{1+2+3+4+5}{5}=\mathbf{3}$

母分散：$\sigma^2=\dfrac{1^2+2^2+3^2+4^2+5^2}{5}-9=\mathbf{2}$

(b)　可能な標本は次に示す ${}_5C_2=10$ 通りである．

標本	(1, 2)	(1, 3)	(1, 4)	(1, 5)	(2, 3)	(2, 4)	(2, 5)	(3, 4)	(3, 5)	(4, 5)
$\bar{x}$	1.5	2	2.5	3	2.5	3	3.5	3.5	4	4.5

(c)　よって，$\overline{X}$ の標本分布は

$\bar{x}$	1.5	2	2.5	3	3.5	4	4.5
$P(\overline{X}=\bar{x})$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{2}{10}$	$\frac{2}{10}$	$\frac{2}{10}$	$\frac{1}{10}$	$\frac{1}{10}$

(d)　$\overline{X}$ の標本分布より

$$E(\overline{X})=\mathbf{3} \quad (\text{分布の対称性より})$$

$$V(\overline{X})=1.5^2\times\frac{1}{10}+2^2\times\frac{1}{10}+\cdots+4.5^2\times\frac{1}{10}-3^2=\mathbf{0.75}$$

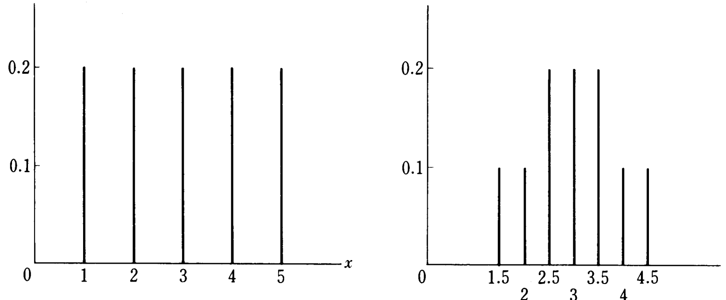
一方，公式からは

$$E(\overline{X})=\mu=3$$

$$V(\overline{X})=\frac{N-n}{N-1}\cdot\frac{\sigma^2}{n}=\frac{5-2}{5-1}\cdot\frac{2}{2}=0.75$$

これらの値は $\overline{X}$ の標本分布から求めた値に一致している．

(e)



母集団分布　　　$\overline{X}$ の標本分布

例題 2　**(平均とメジアンの標本分布)**

1, 2, 3 の目がそれぞれ 2 個ずつ記入されたサイコロを 1 回投げるとき，出る目の平均 μ と分散 σ^2 を求めよ．

(a) このサイコロを 3 回投げるとき，(i) 出る目の平均，(ii) 出る目のメジアン，の標本分布を導け．

(b)　これら標本分布の平均はいずれも μ に等しいことを示し，どちらの分散が小さいかを述べよ．

解　(a)　このサイコロの 1 回の投げの結果を X とすると，X の確率分布は

x	1	2	3
$P(X=x)$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

よって，$\mu=1\times\frac{1}{3}+2\times\frac{1}{3}+3\times\frac{1}{3}=2$

$$\sigma^2=1^2\times\frac{1}{3}+2^2\times\frac{1}{3}+3^2\times\frac{1}{3}-2^2=\frac{2}{3}$$

サイコロの3回の投げの結果を X_1, X_2, X_3，その標本平均を $\overline{X}$，標本メジアンを M とし，$\overline{X}$ と M の標本値をそれぞれ $\bar{x}$ と m とする．

(X_1, X_2, X_3) のすべての可能な結果と，それに対する $\bar{x}$ の値，および m の値を次の表に示す．

可能な結果 (x_1, x_2, x_3)	$\bar{x}$	m	可能な結果 (x_1, x_2, x_3)	$\bar{x}$	m
(1, 1, 1)	1	1	(3, 3, 1)	$\frac{7}{3}$	3
(1, 1, 2)	$\frac{4}{3}$	1	(3, 1, 3)	$\frac{7}{3}$	3
(1, 2, 1)	$\frac{4}{3}$	1	(1, 3, 3)	$\frac{7}{3}$	3
(2, 1, 1)	$\frac{4}{3}$	1	(3, 3, 2)	$\frac{8}{3}$	3
(1, 1, 3)	$\frac{5}{3}$	1	(3, 2, 3)	$\frac{8}{3}$	3
(1, 3, 1)	$\frac{5}{3}$	1	(2, 3, 3)	$\frac{8}{3}$	3
(3, 1, 1)	$\frac{5}{3}$	1	(1, 2, 3)	2	2
(2, 2, 1)	$\frac{5}{3}$	2	(1, 3, 2)	2	2
(2, 1, 2)	$\frac{5}{3}$	2	(2, 1, 3)	2	2
(1, 2, 2)	$\frac{5}{3}$	2	(2, 3, 1)	2	2
(2, 2, 2)	2	2	(3, 1, 2)	2	2
(2, 2, 3)	$\frac{7}{3}$	2	(3, 2, 1)	2	2
(2, 3, 2)	$\frac{7}{3}$	2	(3, 3, 3)	3	3
(3, 2, 2)	$\frac{7}{3}$	2			

これら27通りの結果の得られる確率はいずれも $\frac{1}{3}\times\frac{1}{3}\times\frac{1}{3}=\frac{1}{27}$ である．

よって，この表より $\overline{X}$ の標本分布と M の標本分布は

(i)

$\bar{x}$	1	$\frac{4}{3}$	$\frac{5}{3}$	2	$\frac{7}{3}$	$\frac{8}{3}$	3
$P(\bar{X}=\bar{x})$	$\frac{1}{27}$	$\frac{3}{27}$	$\frac{6}{27}$	$\frac{7}{27}$	$\frac{6}{27}$	$\frac{3}{27}$	$\frac{1}{27}$

(ii)

m	1	2	3
$P(M=m)$	$\frac{7}{27}$	$\frac{13}{27}$	$\frac{7}{27}$

(b)　これら2つの分布はいずれも2を中心として対称であるから

$$E(\overline{X})=E(M)=2=\mu$$

$\overline{X}$ と M の分散は，それぞれ

$$V(\overline{X})=1^2\times\frac{7}{27}+\left(\frac{4}{3}\right)^2\times\frac{3}{27}+\cdots+3^2\times\frac{1}{27}-4=\boldsymbol{\frac{2}{9}}$$

$$V(M)=1^2\times\frac{7}{27}+2^2\times\frac{13}{27}+3^2\times\frac{7}{27}-4=\boldsymbol{\frac{14}{27}}$$

よって，

$$\boldsymbol{V(\overline{X})<V(M)}.$$

例題 3　($\overline{X}$ **の標本分布**)

ある工場で生産される電球の寿命は平均1180時間，標準偏差20時間の正規分布に従う．このとき，次を求めよ．

(a)　25個の電球の無作為標本の寿命の平均が1170時間を超える確率．

(b)　無作為に選んだ n 個の電球の寿命の平均が，少なくとも0.9の確率で1175時間を超えるといえる n の値．

解　電球の寿命を X とすると

$$X\sim N(1180,\ 20^2)$$

25個の電球の寿命を $X_1, X_2, \cdots, X_{25}$，その平均を

$$\overline{X}=\frac{1}{25}\sum_{i=1}^{25}X_i$$

とすれば，

$$E(\overline{X})=1180, \qquad V(\overline{X})=\frac{20^2}{25}=16$$

よって，$\overline{X}\sim N(1180, 4^2)$ であるから

（a）　$P(\overline{X}>1170)=P\left(Z>\dfrac{1170-1180}{4}\right)$

$$=1-\Phi(-2.5)=\Phi(2.5)=\mathbf{0.9938}$$

（b）　n 個の電球の寿命を $X_1, X_2, \cdots, X_n$，その平均を

$$\overline{X}_n=\frac{1}{n}\sum_{i=1}^{n}X_i$$

とすれば，

$$\overline{X}_n\sim N\left(1180, \frac{20^2}{n}\right)$$

与えられた情報より，

$$P(\overline{X}_n>1175)\geqq 0.90$$
$$P(\overline{X}_n<1175)<0.10$$
$$P\left(Z<\frac{1175-1180}{20/\sqrt{n}}\right)\leqq 0.10$$
$$\Phi\left(-\frac{\sqrt{n}}{4}\right)\leqq 0.10, \qquad \Phi\left(\frac{\sqrt{n}}{4}\right)\geqq 0.90$$

正規分布表より

$$\frac{\sqrt{n}}{4}\geqq 1.28 \quad\Rightarrow\quad n\geqq 26.2$$

よって，**27 個以上**.

例題 4　**（正規変量の 1 結合）**

X_1 と X_2 は独立な確率変数で，

$$X_1\sim N(5, 2), \qquad X_2\sim N(3, 1)$$

のとき，$Y=2X_1-X_2$ に対して

（a）　$E(Y)$,

（b）　$V(Y)$,

（c）　$P(Y>8.5)$

を求めよ.

解　（a）　$E(Y)=E(2X_1-X_2)=2E(X_1)-E(X_2)=2\times 5-3=\mathbf{7}$

（b）　$V(Y)=V(2X_1-X_2)$

$$=4V(X_1)+V(X_2) \qquad (X_1 \text{ と } X_2 \text{ は独立であるから})$$
$$=4\times 2+1=\mathbf{9}$$

（c）正規分布に従う独立な確率変数 X_1, X_2 の一次結合 $a_1X_1+a_2X_2$ は正規分布に従うから

$$Y \sim N(7,9)$$

よって,

$$\begin{aligned} P(Y>8.5)&=1-P(Y<8.5)\\ &=1-P\left(Z<\frac{8.5-7}{3}\right)\\ &=1-P(Z<0.5)\\ &=1-\Phi(0.5)=1-0.6915=\mathbf{0.3085} \end{aligned}$$

例題 5　(2 つの平均の差の分布)

確率変数 X と Y は独立で, $X \sim N(2,1)$, $Y \sim N(3,2)$ とする. $\overline{X}$ と $\overline{Y}$ は X と Y に関するそれぞれ 5 個の観測値の平均を表す. このとき, 次を求めよ.

（a）$E[(X-3)^2]$
（b）$E[(X+2Y)^2]$
（c）$\overline{X}-\overline{Y}$ の分布
（d）$P(\overline{X}>\overline{Y})$

解　（a）
$$\begin{aligned} E[(X-3)^2]&=E[(X-2-1)^2]\\ &=E[(X-2)^2-2(X-2)+1]\\ &=E[(X-2)^2]-2E(X-2)+1\\ &=1-0+1=\mathbf{2} \end{aligned}$$

（b）
$$\begin{aligned} E[(X+2Y)^2]&=E[\{X-2+2(Y-3)+8\}^2]\\ &=E[(X-2)^2]+4E[(Y-3)^2]+64\\ &\quad+4E[(X-2)(Y-3)]+16E(X-2)+32E(Y-3) \end{aligned}$$

仮定より

$E[(X-2)^2]=1$,　$E[(Y-3)^2]=2$,

$E(X-2)=0$,　$E(Y-3)=0$,

$$\begin{aligned} E[(X-2)(Y-3)]&=E(X-2)E(Y-3) \qquad (X-2 \text{ と } Y-3 \text{ は独立だから})\\ &=0 \end{aligned}$$

であるから,

$$E[(X+2Y)^2]=1+4\times 2+64=\mathbf{73}$$

（c）　X, Y は正規変数であるから, $\overline{X}-\overline{Y}$ も正規変数である.

$$E(\overline{X}-\overline{Y})=E(\overline{X})-E(\overline{Y})=2-3=-1$$

$$V(\overline{X}-\overline{Y})=V(\overline{X})+V(\overline{Y})$$

$$=\frac{1}{5}+\frac{2}{5}=\frac{3}{5}$$

よって,

$$\overline{X}-\overline{Y}\sim \boldsymbol{N}\left(\mathbf{-1},\ \mathbf{\frac{3}{5}}\right)$$

（d）　$P(\overline{X}>\overline{Y})=P(\overline{X}-\overline{Y}>0)$

$$=P\left(Z>\frac{0+1}{\sqrt{3/5}}\right)=1-\Phi(1.291)=1-0.9017=\mathbf{0.0983}$$

例題 6　**（正規変量の和）**

大型のりんご 1 個の重さは平均 330 g, 標準偏差 15 g の正規分布に従い, 並型のりんご 1 個の重さは平均 280 g, 標準偏差 10 g の正規分布に従う. そのとき, 次の確率を求めよ.

（a）　無作為に選んだ大型のりんご 3 個の重さの合計が 1000 g を超える.

（b）　大型のりんご 1 個と並型のりんご 1 個を無作為に選ぶとき, 並型の重さが大型の重さを超える.

（c）　無作為に選んだ並型のりんご 5 個の重さの合計が, 無作為に選んだ大型のりんご 4 個の重さの合計を超える.

解　大型のりんご 1 個の重さを X, 並型のりんご 1 個の重さを Y とすると

$$X\sim N(330,\ 15^2),\quad Y\sim N(280,\ 10^2)$$

（a）　大型のりんご 3 個の重さを X_1, X_2, X_3 とすれば, ,$T=X_1+X_2+X_3$ の分布は

$$E(T)=E(X_1+X_2+X_3)$$
$$=E(X_1)+E(X_2)+E(X_3)$$
$$=330+330+330=990$$

$$V(T)=V(X_1+X_2+X_3)$$
$$=V(X_1)+V(X_2)+V(X_3)\qquad (X_1, X_2, X_3 \text{ は独立だから})$$