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	

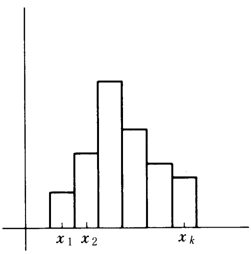
1-2　グラフによる表現

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．

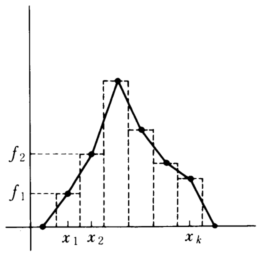
度数多角形　ヒストグラムの各長方形の上辺の中点を結んで得られる折れ線グラフ．

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．

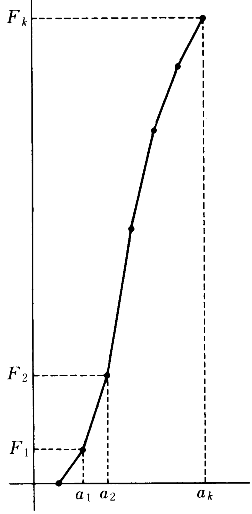
ヒストグラム



度数多角形



累積度数多角形



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

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

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

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

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

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

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

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

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)

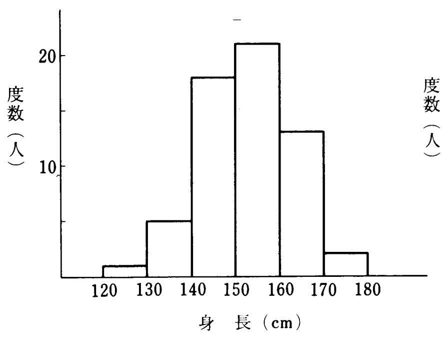
次の度数分布において，階級の真の限界，階級値および相対度数を与えよ．また，ヒストグラム，度数多角形，および累積度数多角形を図示せよ．

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

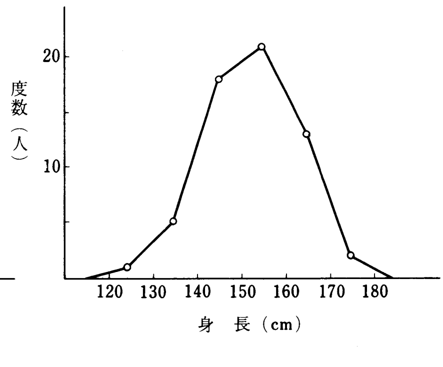
解

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，講義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.

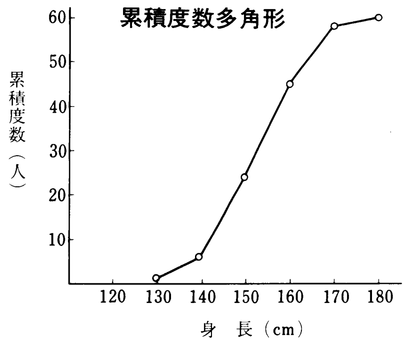
本書は各章の始めに公式や定義などを要項の形で簡潔にまとめ，その後に例題と演習問題が続く. 例題は理論的に難しい問題は避け，標準的でかつ興味ある実際的な問題を慎重に選び，すべてにていねいな解答を与えた. また，章末の演習問題も例題と同種なもので，これにも詳しい解答をつけている. 文系，理系を問わず広く学生諸君らが容易に理解できるよう，数学の予備知識としては高校の数学Ｉ，基礎解析の程度とした.

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	正正𠃊	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	正𠃊	174.5-179.5	177	9
180-184	𠃊	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，

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

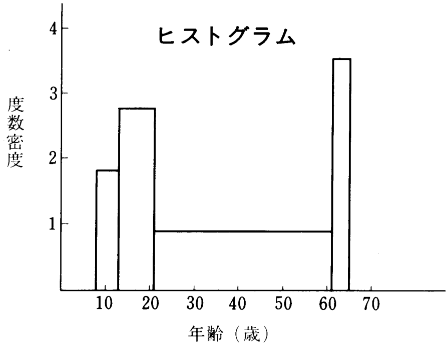
の関係より，

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

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	正丅	7	36
5	一	1	37
6	下	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

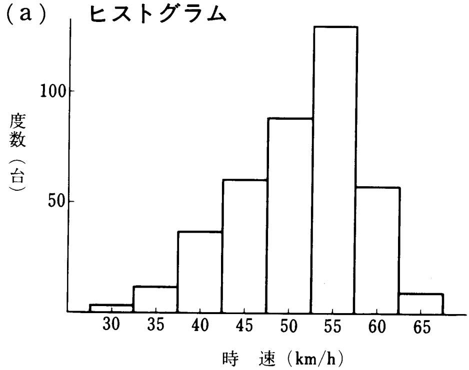
この結果を

（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.

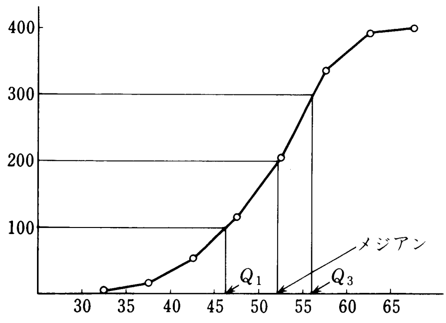
解　(a)　**ヒストグラム**



(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：

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

3 年生全員の身長の平均は

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

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

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

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 g,　250 g,　239 g,　264 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 $\dfrac{r}{n}$ 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**

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

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　$$P(E_k|B) = \frac{P(E_k)P(B|E_k)}{\sum_{j=1}^{n} P(E_j)P(B|E_j)} \qquad (k=1, 2, \cdots, n)$$

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．

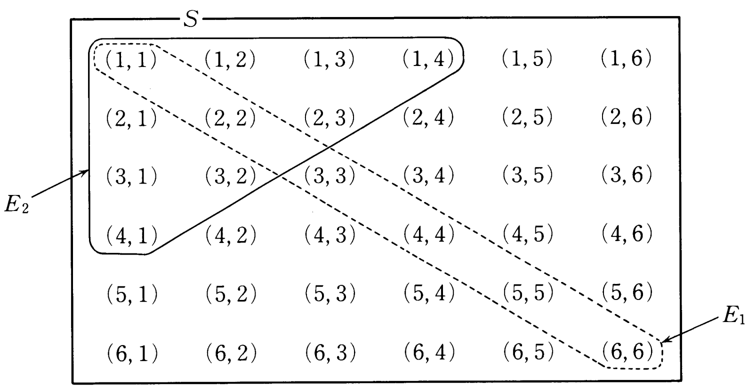
（a）　標本空間 S.

（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,

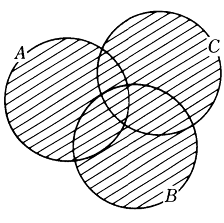
（c） A 紙のみ

を購読する世帯の割合を求めよ．

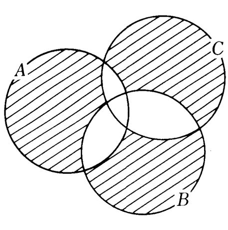
解 ベン図を使って解く．題意から

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

（a）
$$\begin{aligned}P(A\cup B\cup C)&=P(A)+P(B)+P(C)-P(A\cap B)-P(A\cap C)\\&\qquad -P(B\cap C)+P(A\cap B\cap C)\\&=0.20+0.16+0.14-0.08-0.05-0.04+0.02\\&=\mathbf{0.35}\end{aligned}$$



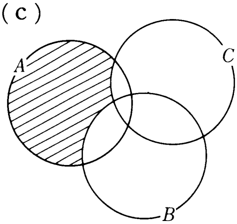
（b）



左図より

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

（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）

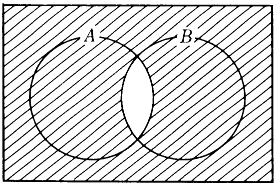
$P(A)=a$, $P(B)=b$, $P(A\cap B)=c$ とするとき，次の事象の確率を a, b, c で表せ．

（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）

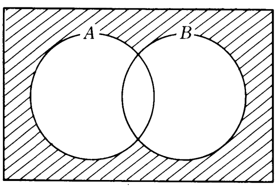


$P(\overline{A}\cup\overline{B})=P(\overline{A\cap B})$

$=1-P(A\cap B)$

$=\mathbf{1-c}$

（b）



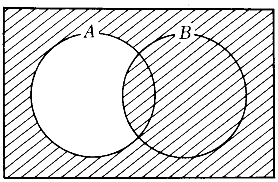
$P(\overline{A}\cap\overline{B})=P(\overline{A\cup B})$

$=1-P(A\cup B)$

$=1-\{P(A)+P(B)-P(A\cap B)\}$

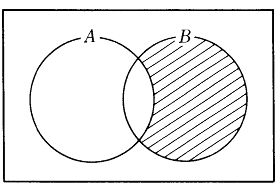
$=\mathbf{1-a-b+c}$

（c）



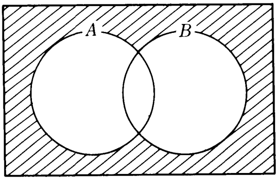
$P(\overline{A}\cup B)=\mathbf{1-a+c}$

（d）



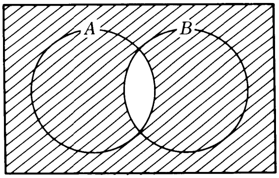
$P(\overline{A}\cap B)=\mathbf{b-c}$

（e）



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

（f）



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

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

$$P(A\cap B)=P(A)P(B|A)$$
$$=0.5\times0.9=\mathbf{0.45}$$

（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

$$\frac{{}_{4}C_{4}\times{}_{48}C_{9}}{{}_{53}C_{13}}=\frac{\dfrac{48!}{9!39!}}{\dfrac{52!}{13!39!}}=\mathbf{0.0026}$$

（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

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

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

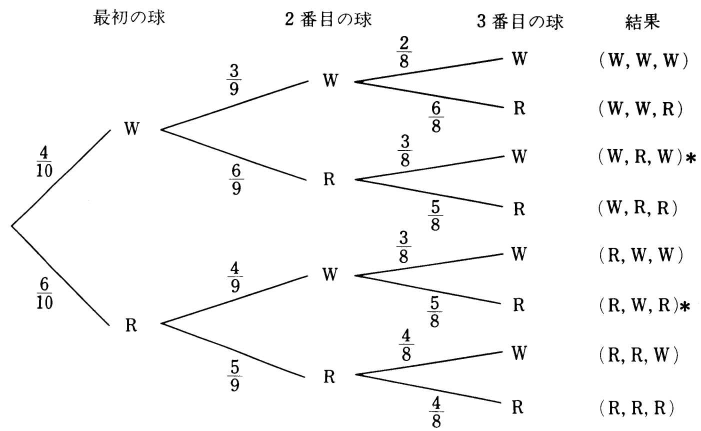
$$(0.9974)^n \leqq 0.5$$
$$n\log 0.9974 \leqq \log 0.5$$
$$n \geqq 266.24$$

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)　$F(-\infty)=0,\quad F(+\infty)=1$

(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

$$E(X)=30{,}000{,}000\times\frac{10}{10{,}000{,}000}+150{,}000\times\frac{90}{10{,}000{,}000}+\cdots$$
$$+300\times\frac{2{,}000{,}000}{10{,}000{,}000}+0\times\frac{7{,}998{,}300}{10{,}000{,}000}=\mathbf{139.52}\ (円)$$

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) $V(2X+3Y)=4V(X)+9V(Y)=4\sigma^2+9\times2\sigma^2=\boldsymbol{22\sigma^2}$

(e) $V(X-3Y-5)=V(X)+9V(Y)=\sigma^2+9\times2\sigma^2=\boldsymbol{19\sigma^2}$

(f) $V(aX-bY)=a^2V(X)+b^2V(Y)=a^2\sigma^2+b^2\times2\sigma^2=\boldsymbol{(a^2+2b^2)\sigma^2}$

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	$\frac{3}{7}$	$\frac{4}{7}\cdot\frac{3}{6}$	$\frac{4}{7}\cdot\frac{3}{6}\cdot\frac{3}{5}$	$\frac{4}{7}\cdot\frac{3}{6}\cdot\frac{2}{5}\cdot\frac{3}{4}$	$\frac{4}{7}\cdot\frac{3}{6}\cdot\frac{2}{5}\cdot\frac{1}{4}\cdot1$
$P(X=x)$	$\frac{3}{7}$	$\frac{2}{7}$	$\frac{6}{35}$	$\frac{3}{35}$	$\frac{1}{35}$

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，

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

$$V(X)=E(X^2)-2^2$$

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

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
$P(X=x)$	$\dfrac{2}{8}$	$\dfrac{3}{8}$	$\dfrac{3}{8}$

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，

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

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

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

$E(Z)=\mathbf{3}$　（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

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

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) $\displaystyle\int_1^2 \frac{a}{x^3}\,dx=1$ 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，$y=\sqrt{\mathbf{2}}$.

（c）　メジアンを M とすると

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

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

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

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）

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

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

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

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

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

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

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

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

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

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

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

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

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

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

例題 14　（分布関数と密度関数）

X の累積分布関数が

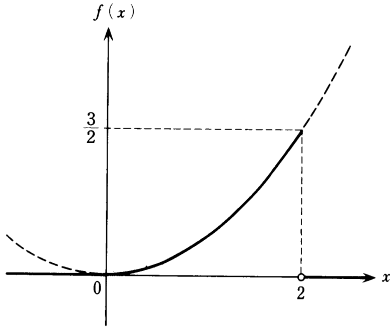
$$F(x)=\begin{cases} 0 & (x<0) \\ kx^3 & (0<x\leqq 2) \\ 1 & (x>2) \end{cases}$$

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

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

$f(x)$ 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

$x \backslash y$	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

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

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

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

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

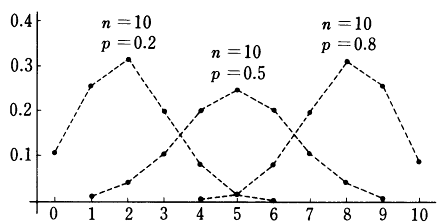
2項分布　1回の試行である事象の起こる確率を p とする．n 回の独立試行でこの事象の起こる回数 X の確率分布は

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

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

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

（ii）$(n+1)p$ が整数でないならば，モードは

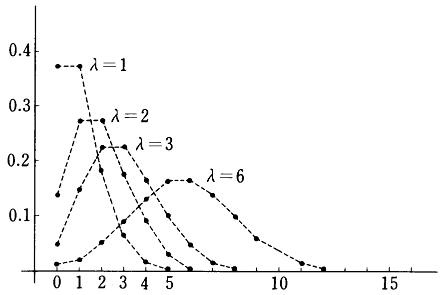
$$(n+1)p \text{ を超えない最大の整数値.}$$

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

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

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

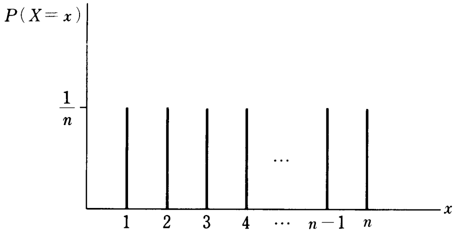
$$P(X=x)=\begin{cases} \dfrac{1}{n} & (x=1, 2, \cdots, n) \\ 0 & (\text{その他}) \end{cases}$$

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

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

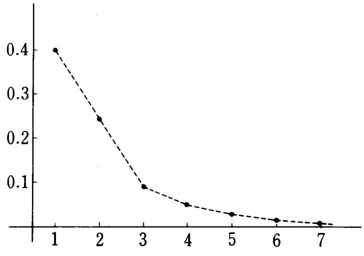
で与えられるとき，

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

$$E(X)=np=8\times0.8=\mathbf{6.4}$$
$$V(X)=npq=8\times0.8\times0.2=\mathbf{1.28}$$

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

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

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，

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

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

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 $B\left(5, \frac{1}{3}\right)$ 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 $\frac{1}{2}$ 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 $B\left(8, \frac{1}{2}\right)$ 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）．

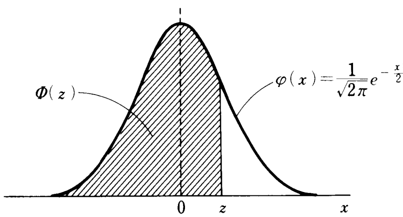


図 1

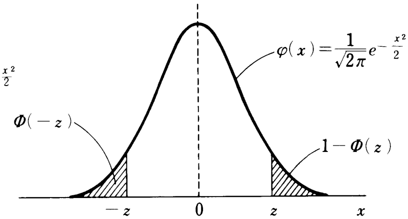


図 2

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

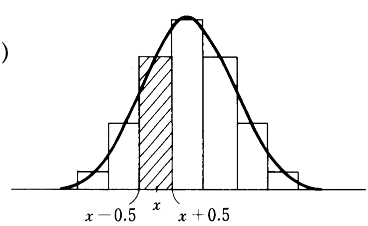
$$np>5 \text{ かつ } nq>5$$

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．

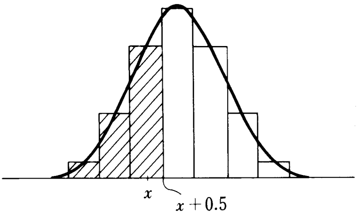
5-2　半整数補正

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，

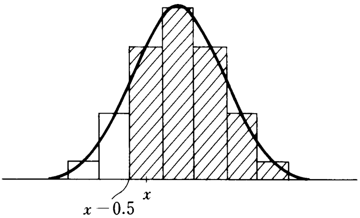
$$P(X=x)\cong P(x-0.5<Y<x+0.5)$$



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

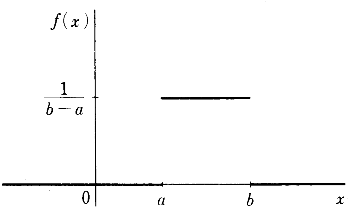


$$P(X \geqq x) \cong P(Y > x-0.5)$$



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**(くけい)**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

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

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）　$P(X>8)$

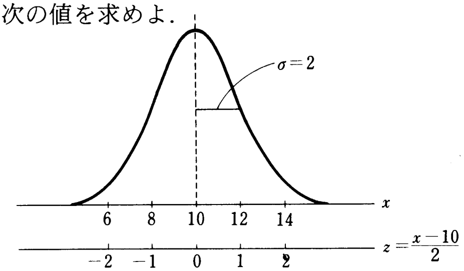
（d）　$P(X<7)$

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

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

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

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

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

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

$$=P(0.2<Z<1.1)\qquad \text{（分布の対称性より）}$$

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

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

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

（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，

$$\frac{c-10}{2}=1.03+\frac{0.85-0.8485}{0.8508-0.8485}\times0.01\fallingdotseq1.037$$

$$c=10+2\times1.037\fallingdotseq\mathbf{12.07}$$

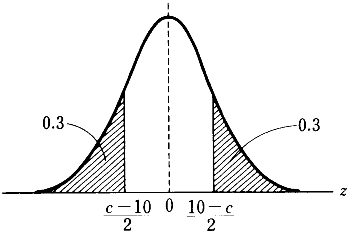
（h）　$P(X<c)=\Phi\left(\dfrac{c-10}{2}\right)$ より，c は

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

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

$$\frac{10-c}{2}=0.52+\frac{0.7-0.6985}{0.7019-0.6985}\times0.01\fallingdotseq0.5234$$

$$c=10-2\times0.5234\fallingdotseq\mathbf{8.953}$$

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)　$P(|X|<0.3)$

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

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

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

$$=\Phi(1.5)-\Phi(-1.5)$$
$$=2\Phi(1.5)-1=2\times0.9332-1=\mathbf{0.8664}$$

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

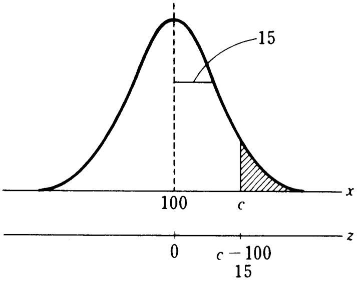
（a） $P(X<85)=P\left(Z<\dfrac{85-100}{15}\right)$
$$=\Phi(-1)=1-\Phi(1)=1-0.8413=\mathbf{0.1587}$$

（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

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

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，

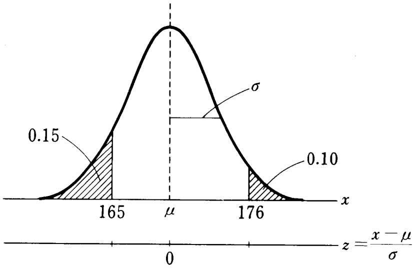
$$X \sim N(\mu, \sigma^2)$$

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

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

(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．

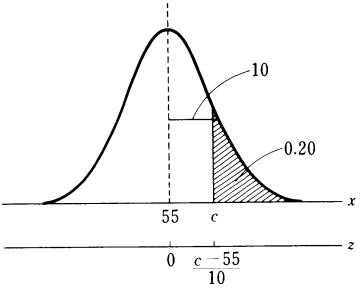
解　得点を X とすると，$X \sim N(55, 10^2)$.

よって

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

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

$300\times0.2417=72.51$. 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

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

よって，補間により

$$\frac{c-55}{10}=0.84+\frac{0.8-0.7995}{0.8023-0.7995}\times0.01\fallingdotseq0.842$$

$$c=55+10\times0.842=63.42$$

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） $P\{(8:50<T<9:55)\cap(X\geqq 3)\}$

$=P(8:50<T<9:55)P(X\geqq 3)$ 　（T と X 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 $\dfrac{1}{3}$ 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

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

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

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

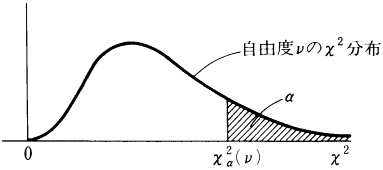
で与えられる分布を**自由度 ν の χ^2 分布**という. ここで, 定数 c は $f_\nu(x)$ が密度関数であるという条件から決まる.

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

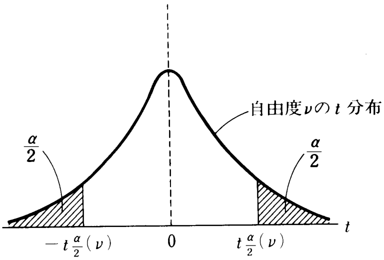
$$g_\nu(x)=c\left(1+\frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}\qquad(-\infty<x<\infty)$$

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，

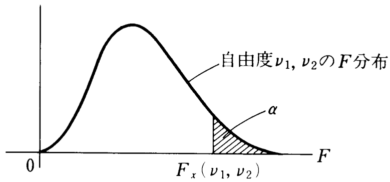
$$s^2=\frac{n_1{s_1}^2+n_2{s_2}^2}{n_1+n_2-2}$$

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 (ν_1, ν_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

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

$$V(\overline{X})=1.5^2\times\frac{1}{10}+2^2\times\frac{1}{10}+\cdots+4.5^2\times\frac{1}{10}-3^2=\mathbf{0.75}$$

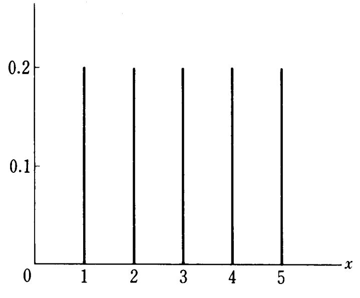
一方，公式からは

$$E(\overline{X})=\mu=3$$

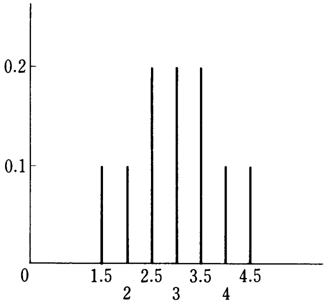
$$V(\overline{X})=\frac{N-n}{N-1}\cdot\frac{\sigma^2}{n}=\frac{5-2}{5-1}\cdot\frac{2}{2}=0.75$$

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）　$V(Y)=V(2X_1-X_2)$

$$=4V(X_1)+V(X_2) \quad (X_1 \text{ と } X_2 \text{ は独立であるから})$$
$$=4\times 2+1=\mathbf{9}$$

（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

$E[(X-2)^2]=1$, $\quad E[(Y-3)^2]=2$,

$E(X-2)=0$, $\quad E(Y-3)=0$,

$$\begin{aligned}E[(X-2)(Y-3)]&=E(X-2)E(Y-3) \quad (X-2 \text{ と } Y-3 \text{ は独立だから})\\&=0\end{aligned}$$

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,

$$\overline{X}-\overline{Y}\sim N\left(\mathbf{-1},\ \mathbf{\frac{3}{5}}\right)$$

（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

$$E(T)=E(X_1+X_2+X_3)$$

$$=E(X_1)+E(X_2)+E(X_3)$$

$$=330+330+330=990$$

$$V(T)=V(X_1+X_2+X_3)$$

$$=V(X_1)+V(X_2)+V(X_3)\qquad (X_1,\ X_2,\ X_3 \text{ は独立だから})$$

$$=15^2+15^2+15^2=675$$

より,

$$T \sim N(990, 675)$$

よって,

$$P(T>1000)=P\left(Z>\frac{1000-990}{\sqrt{675}}\right)$$
$$\fallingdotseq 1-\Phi(0.385)=1-0.6499=\mathbf{0.3501}$$

（b）　$X-Y$ の分布は

$$E(X-Y)=E(X)-E(Y)$$
$$=330-280=50$$
$$V(X-Y)=V(X)+V(Y)$$
$$=15^2+10^2=325$$

より,

$$X-Y \sim N(50, 325)$$

よって,

$$P(Y>X)=P(X-Y<0)=P\left(Z<\frac{0-50}{\sqrt{325}}\right)$$
$$=1-\Phi(2.774)=1-0.9972=\mathbf{0.0028}$$

（c）　大型りんご 4 個の重さを X_1, X_2, X_3, X_4, その合計を $T_1=\sum_{i=1}^{4} X_i$,

並型りんご 5 個の重さを Y_1, Y_2, Y_3, Y_4, Y_5 その合計を $T_2=\sum_{i=1}^{5} Y_i$

とすれば, 仮定より

$$T_1 \sim N(1320, 4\times 15^2)$$
$$T_2 \sim N(1400, 5\times 10^2)$$

よって,

$$T_2-T_1 \sim N(80, 5\times 10^2+4\times 15^2)=N(80, 1400)$$

ゆえに,

$$P(T_2>T_1)=P(T_2-T_1>0)$$
$$=P\left(Z>\frac{0-80}{\sqrt{1400}}\right)$$
$$=1-\Phi(2.138)=1-0.9837=\mathbf{0.0163}$$

例題 7（$\overline{X}$ の平均と分散）

$X_1, X_2, \cdots, X_n$ を平均 μ，分散 σ^2 の母集団からの大きさ n の無作為標本とし，その平均を $\overline{X}$ とするとき

$$E(\overline{X})=\mu, \qquad V(\overline{X})=\frac{\sigma^2}{n}$$

を示せ．

解　$X_1, X_2, \cdots, X_n$ は同一母集団からの無作為標本であるから，互いに独立で，かつ

$$E(X_1)=E(X_2)=\cdots=E(X_n)=\mu$$
$$V(X_1)=V(X_2)=\cdots=V(X_n)=\sigma^2$$

よって，

$$\begin{aligned}E(\overline{X})&=E\left(\frac{X_1+X_2+\cdots+X_n}{n}\right)\\&=\frac{1}{n}E(X_1+X_2+\cdots+X_n)\\&=\frac{1}{n}\{E(X_1)+E(X_2)+\cdots+E(X_n)\}\\&=\frac{1}{n}(\mu+\mu+\cdots+\mu)\\&=\mu\end{aligned}$$

$$\begin{aligned}V(\overline{X})&=V\left(\frac{X_1+X_2+\cdots+X_n}{n}\right)\\&=\frac{1}{n^2}V(X_1+X_2+\cdots+X_n) \qquad (X_1, X_2, \cdots, X_n \text{ の独立性より})\\&=\frac{1}{n^2}\{V(X_1)+V(X_2)+\cdots+V(X_n)\}\\&=\frac{1}{n^2}(\sigma^2+\sigma^2+\cdots+\sigma^2)\\&=\frac{\sigma^2}{n}\end{aligned}$$

例題 8　(独立な正規変数の和と差)

ある大学の男子学生の体重は平均 60 kg, 標準偏差 8 kg の正規分布に従い, 女子学生の体重は平均 52 kg, 標準偏差 5 kg の正規分布に従うことが知られている. いま, 1 人の男子学生と 1 人の女子学生を無作為に選ぶとき, 次の確率を求めよ.

(a)　2 人の体重の合計が 100 kg を超える.

(b)　女子学生の方が男子学生よりも重い.

(c)　女子学生の体重が男子学生の体重の少なくとも $\dfrac{3}{4}$ はある.

解　男子学生の体重を X, 女子学生の体重を Y とすると,

$$X \sim N(60, 8^2), \qquad Y \sim N(52, 5^2)$$

(a)　$S=X+Y$ とおくと,

$$E(S)=E(X)+E(Y)=60+52=112$$
$$V(S)=V(X)+V(Y)=8^2+5^2=89$$

よって,

$$S \sim N(112, 89)$$

ゆえに

$$\begin{aligned} P(S>100)&=1-P(S<100)\\ &=1-P\left(Z<\frac{100-112}{\sqrt{89}}\right)\\ &=1-\Phi(-1.272)=\Phi(1.272)=\mathbf{0.8983} \end{aligned}$$

(b)　$D=X-Y$ とおくと,

$$E(D)=E(X)-E(Y)=60-52=8$$
$$V(D)=V(X)+V(Y)=89$$

よって,

$$D \sim N(8, 89)$$

ゆえに

$$\begin{aligned} P(Y>X)&=P(X-Y<0)\\ &=P\left(Z<\frac{0-8}{\sqrt{89}}\right)\\ &=\Phi(-0.848)\\ &=1-\Phi(0.848)=1-0.8017=\mathbf{0.1983} \end{aligned}$$

（c）　$W=Y-\frac{3}{4}X$ とおくと

$$E(W)=E(Y)-\frac{3}{4}E(X)=52-\frac{3}{4}\times 60=7$$

$$V(W)=V(Y)+\left(\frac{3}{4}\right)^2 V(X)=25+\frac{9}{16}\times 64=61$$

よって,

$$P\left(Y>\frac{3}{4}X\right)=P(W>0)$$
$$=P\left(Z>\frac{0-7}{\sqrt{61}}\right)$$
$$=\Phi(0.896)=\mathbf{0.8149}$$

例題 9　**（独立な正規変数の和）**

毎日定刻に自宅を出て会社に通勤しているあるサラリーマンの通勤所要時間（分）は次の 3 段階からなる.

X : 自宅から乗車駅までの歩行時間と下車駅から会社までの歩行時間の合計.

Y : 乗車駅での電車の待ち時間.

Z : 電車に乗っている時間.

X, Y, Z の平均と分散が

	平均	標準偏差
X	20	1
Y	5	3
Z	40	2

で与えられるとき, 次を求めよ.

（a）　このサラリーマンの通勤所要時間の平均と標準偏差.

（b）　X, Y, Z はいずれも正規分布に従うとして, このサラリーマンが自宅を出て 1 時間以内に会社に着く確率.

解　通勤所要時間を W とすると

$$W=X+Y+Z$$

ゆえに,

（a）　$E(W)=E(X+Y+Z)$
$$=E(X)+E(Y)+E(Z)$$
$$=20+5+40=\mathbf{65}\ (分)$$

$$V(W)=V(X+Y+Z)$$
$$=V(X)+V(Y)+V(Z) \qquad (X,\ Y,\ Z\ は互いに独立とみなして)$$
$$=1^2+3^2+2^2=14$$

よって，

$$\sqrt{V(W)}=\sqrt{14}=\mathbf{3.74}\ (分)$$

（b）　W は平均 65，標準偏差 3.74 の正規分布に従うから

$$P(W<60)=P\left(Z<\frac{60-65}{\sqrt{14}}\right)=1-\Phi(1.336)\fallingdotseq\mathbf{0.0908}$$

例題 10　（中心極限定理）

ある種の繊維 1 本の破断強度は平均 1.2 kg，標準偏差 0.1 kg の正規分布に従う．この繊維 100 本の束に 122.5 kg の荷重をかけるとき，束が破壊される確率を求めよ．

解　繊維 100 本の破断強度を $X_1, X_2, \cdots, X_{100}$ とすると

$$X_i \sim N(1.2, 0.1^2) \qquad (i=1, 2, \cdots, 100)$$

100 本の束の破断強度を Y とすると

$$Y=X_1+X_2+\cdots+X_{100}$$

よって，Y は

$$E(Y)=E(X_1)+E(X_2)+\cdots+E(X_{100})=100\times1.2=120$$
$$V(Y)=V(X_1)+V(X_2)+\cdots+V(X_{100})=100\times0.1^2=1$$

の正規分布に従うから，

$$P(Y>122.5)=P\left(Z>\frac{122.5-120}{1}\right)=1-\Phi(2.5)=\mathbf{0.006}$$

［注］　この問題では，$n=100$ より中心極限定理が使えるので，繊維の破断強度が**正規分布**に従うとの仮定がなくても解ける．

例題 11　（中心極限定理）

乱数表から 1 桁の乱数を 50 個とるとき，その平均が 4 と 5 の間にある確率を求めよ．

解　乱数の分布は離散型一様分布

x	0	1	2	3	4	5	6	7	8	9
$P(X=x)$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$

で，この分布の平均と分散は

$$\mu=0\times\frac{1}{10}+1\times\frac{1}{10}+\cdots+9\times\frac{1}{10}=4.5$$

$$\sigma^2=0^2\times\frac{1}{10}+1^2\times\frac{1}{10}+\cdots+9^2\times\frac{1}{10}-4.5^2=8.25$$

である．とられた50個の乱数はこの分布をもつ母集団からの大きさ50の無作為標本とみなされるから，50個の標本の平均を $\overline{X}$ とすれば

$$E(\overline{X})=\mu=4.5$$

$$V(\overline{X})=\frac{\sigma^2}{n}=\frac{8.25}{50}=0.165$$

$n=50$ は大きいから，中心極限定理によって，$\overline{X}$ は近似的に $N(4.5, 0.165)$ に従う．よって，

$$P(4<\overline{X}<5)=P\left(\frac{4-4.5}{\sqrt{0.165}}<Z<\frac{5-4.5}{\sqrt{0.165}}\right)$$

$$=2\Phi(1.231)-1=2\times0.8909-1=\mathbf{0.7818}$$

6章の問題

6.1　サイコロを2回投げるとき，出た目の和の標本分布を求め，この分布の平均と分散を求めよ．

6.2　1の目を3つ，2の目を2つ，3の目を1つもつサイコロを1回投げるとき，(a) 出る目 X の確率分布を示し，(b) その平均と分散を求めよ．(c) このサイコロを7回投げるとき，出る目の平均 $\overline{X}$ の平均と分散を求めよ．

6.3　X と Y は独立な確率変数で，

$$X\sim N(5,9),\qquad Y\sim N(7,16)$$

のとき，

(a)　$X-Y$,

（b） $3X+Y$,

（c） $\overline{X}$,

（d） $\overline{X}-\overline{Y}$

の分布を求めよ．ただし，$\overline{X}$ は $N(5,9)$ からとった大きさ 25 個の無作為標本の平均で，$\overline{Y}$ は $N(7,16)$ からとった大きさ 16 個の無作為標本の平均である．

6.4 ある中学校の男子 3 年生の身長の分布は近似的に平均 163.0 cm，標準偏差 8 cm の正規分布に従い，女子 3 年生の身長の分布は近似的に平均 155.5 cm，標準偏差 6 cm の正規分布に従う．そのとき，次の確率を求めよ．

（a） 男子生徒 2 人を無作為に選ぶとき，2 人の身長の差が 5 cm を超える．

（b） 女子生徒 2 人を無作為に選ぶとき，2 人の身長の差が 5 cm を超える．

（c） 男子生徒 1 人と女子生徒 1 人を無作為に通ぶとき，2 人の身長の差が 10 cm を超える．

6.5 高校 3 年生の身長は次のような正規分布に従うことが知られている．

男子： 169.3 cm，標準偏差 4.6 cm

女子： 156.6 cm，標準偏差 4.2 cm

これら 2 つの正規母集団からとった大きさ 64 と 36 の標本平均をそれぞれ $\overline{X}$, $\overline{Y}$ とするとき，次の値を求めよ．

（a） $\overline{X}$, $\overline{Y}$ の平均と標準偏差．

（b） $\overline{X}-\overline{Y}$ の平均と標準偏差．

（c） $\overline{X}$ が $\overline{Y}$ を少なくとも 12 cm 超える確率．

6.6 桃の缶詰の中味の重さは平均 300 g，標準偏差 1.2 g の正規分布に従い，空き缶の重さは平均 20 g，標準偏差 0.5 g の正規分布に従う．このとき，次の値を求めよ．

（a） 桃の缶詰 1 個の重さの平均と標準偏差．

（b） 無作為に選んだ桃の缶詰 1 個の重さが 323 g を超える確率．

6.7 鋼板にあけた穴にボルトをはめこむ．ボルトの直径は平均 2.80 cm，標準偏差 0.03 cm の正規分布に従い，穴の直径は平均 2.90 cm，標準偏差 0.04 cm の正規分布に従うとき，次を求めよ．

（a） ボルトと穴をそれぞれ無作為に選んではめるとき，ボルトが穴にはまらない確率.

（b） 無作為にとった5個のボルトがすべて穴にはまる確率.

6.8 正規母集団 $N(\mu, \sigma^2)$ からとられた大きさ n の無作為標本を $X_1, X_2, \cdots, X_n$ とするとき，

$$Y = X_1 + 2X_2 + 3X_3 + \cdots + nX_n$$

はどんな分布に従うか.

6.9 （a） 平均 μ，分散 σ^2 の正規母集団からの大きさ4の無作為標本の平均を $\overline{X}$ とするとき，確率

$$P(|\overline{X} - \mu| < \sigma)$$

の値を求めよ.

（b） 平均80，標準偏差10の母集団から大きさ64の無作為標本をとるとき，標本平均 $\overline{X}$ が82を超える確率を求めよ.

6.10 密度関数

$$f(x) = 6x(1-x) \qquad (0 \leqq x \leqq 1)$$

を分布にもつ母集団から大きさ45の標本がとられた．標本平均を $\overline{X}$ とするとき，次を求めよ.

（a） $\overline{X}$ の標本分布.

（b） $P(\overline{X} < 0.55)$.

（c） $P(\overline{X} > 0.46)$.

6.11 小数点以下を四捨五入することで，測定値をそれに最も近い整数値に丸めるとき，丸めの誤差 X は区間 $\left(-\dfrac{1}{2}, \dfrac{1}{2}\right)$ で一様分布に従う確率変数とみなされる．n 個の測定値をそれぞれそれに最も近い整数値に丸めるとき，これら丸め誤差の平均 $\overline{X}$ の平均と分散を求めよ．n は十分大きいと仮定して，$\overline{X}$ の絶対値が $\dfrac{1}{2\sqrt{n}}$ を超えない確率を求めよ.

7

推　定

7-1　点 推 定

母集団の未知母数 θ を推定するために，この母集団からの大きさ n の標本 $X_1, X_2, \cdots, X_n$ のある関数 $T(X_1, X_2, \cdots, X_n)$ を作り，標本の実現値 $x_1, x_2, \cdots, x_n$ から求めた $T(x_1, x_2, \cdots, x_n)$ で θ の値を推定することを**点推定**という．このとき，統計量 $T(X_1, X_2, \cdots, X_n)$ を**推定量**，その実現値 $T(x_1, x_2, \cdots, x_n)$ を**推定値**という．推定量 T としては，次の性質をもつものが望ましい．

不偏推定量　　$T(X_1, X_2, \cdots, X_n)$ の期待値が θ に等しいとき，すなわち

$$E(T)=\theta$$

のとき，$T(X_1, X_2, \cdots, X_n)$ を θ の**不偏推定量**といい，その実現値 $T(x_1, x_2, \cdots, x_n)$ を θ の**不偏推定値**という．

一致推定量　　任意の正数 ε に対して

$$\lim_{n\to\infty} P(|T-\theta|>\varepsilon)=0$$

のとき，この $T(X_1, X_2, \cdots, X_n)$ を θ の**一致推定量**という．

有効推定量　　2 つの不偏推定量 T_1，T_2 に対して，それぞれの分散が

$$V(T_1)<V(T_2)$$

のとき，T_1 は T_2 に比べて**より有効な推定量**であるという．

すべての不偏推定量のなかで分散が最小な推定量を**有効推定量**という．

7-2　最尤推定量

最尤推定量　　実際に推定量を見つけ出す方法として**最尤推定法**がある．これは母集団の分布を $f(x\,;\theta)$ とするとき，この母集団から標本値 $x_1, x_2, \cdots,$

x_n が得られる確率

$$L(\theta)=f(x_1\,;\theta)f(x_2\,;\theta)\cdots f(x_n\,;\theta)=\prod_{i=1}^{n} f(x_i\,;\theta)$$

を考え，この $L(\theta)$ を最大にする θ の値 $\hat{\theta}(x_1, x_2, \cdots, x_n)$ を θ の推定量とするもので，$L(\theta)$ を標本の**尤度関数**，$\hat{\theta}(X_1, X_2, \cdots, X_n)$ を θ の**最尤推定量**という．

7-3 区間推定

信頼区間　未知母数 θ を 1 つの推定量で推定するのではなく，2 つの統計量 T_1，T_2 を作り，一定の確率で θ が区間 (T_1, T_2) 内に含まれるというように，区間によって推定する方法を**区間推定**という．すなわち，θ が

$$P(T_1<\theta<T_2)=1-\alpha$$

のように表されるとき，区間 $(T_1(x_1, x_2, \cdots, x_n),\ T_2(x_1, x_2, \cdots, x_n))$ を θ の $100(1-\alpha)\%$ **信頼区間**といい，$100(1-\alpha)\%$ をこの区間の**信頼度**（または**信頼係数**）という．また，端点の T_1，T_2 を**信頼限界**という．よく使われるものに，90%，95%，99%信頼区間がある．

7-4 平均，分散および比率の信頼区間

平均 μ の信頼区間

（a）　分散 σ^2 が既知の場合.

μ の $100(1-\alpha)\%$信頼区間：　$\bar{x}-z_{\frac{\alpha}{2}}\sqrt{\dfrac{\sigma^2}{n}}<\mu<\bar{x}+z_{\frac{\alpha}{2}}\sqrt{\dfrac{\sigma^2}{n}}$

（b）　分散 σ^2 が未知の場合.

（i）　μ の $100(1-\alpha)\%$ 信頼区間(大標本)：

$$\bar{x}-z_{\frac{\alpha}{2}}\sqrt{\frac{u^2}{n}}<\mu<\bar{x}+z_{\frac{\alpha}{2}}\sqrt{\frac{u^2}{n}}$$

（ii）　μ の $100(1-\alpha)\%$信頼区間(小標本)：

$$\bar{x}-t_{\frac{\alpha}{2}}(n-1)\sqrt{\frac{u^2}{n}}<\mu<\bar{x}+t_{\frac{\alpha}{2}}(n-1)\sqrt{\frac{u^2}{n}}$$

ここで，u^2 は標本不偏分数で，$u^2=\dfrac{1}{n-1}\sum(x_i-\bar{x})^2$.

2 つの平均の差の信頼区間

（a）　分散 ${\sigma_1}^2$，${\sigma_2}^2$ が既知の場合

$\mu_1-\mu_2$ の $100(1-\alpha)$%信頼区間：

$$\bar{x}_1-\bar{x}_2-z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1{}^2}{n_1}+\frac{\sigma_2{}^2}{n_2}}<\mu_1-\mu_2<\bar{x}_1-\bar{x}_2+z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1{}^2}{n_1}+\frac{\sigma_2{}^2}{n_2}}$$

（b）　分散 $\sigma_1{}^2$, $\sigma_2{}^2$ が未知の場合．

（i）　$\mu_1-\mu_2$ の $100(1-\alpha)$% 信頼区間(大標本)：

$$\bar{x}_1-\bar{x}_2-z_{\frac{\alpha}{2}}\sqrt{\frac{u_1{}^2}{n_1}+\frac{u_2{}^2}{n_2}}<\mu_1-\mu_2<\bar{x}_1-\bar{x}_2+z_{\frac{\alpha}{2}}\sqrt{\frac{u_1{}^2}{n_1}+\frac{u_2{}^2}{n_2}}$$

（ii）　$\mu_1-\mu_2$ の $100(1-\alpha)$%信頼区間(小標本；$\sigma_1{}^2=\sigma_2{}^2$)：

$$\bar{x}_1-\bar{x}_2-t_{\frac{\alpha}{2}}(n_1+n_2-2)\sqrt{\frac{(n_1-1)u_1{}^2+(n_2-1)u_2{}^2}{n_1+n_2-2}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}<\mu_1-\mu_2$$

$$<\bar{x}_1-\bar{x}_2+t_{\frac{\alpha}{2}}(n_1+n_2-2)\sqrt{\frac{(n_1-1)u_1{}^2+(n_2-1)u_2{}^2}{n_1+n_2-2}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}$$

ここで，$u_1{}^2$, $u_2{}^2$ は $\sigma_1{}^2$, $\sigma_2{}^2$ の標本不偏分散．

［注］　大標本とは標本の大きさが約 30 以上のときをいう．

分散 $\boldsymbol{\sigma^2}$ の信頼区間

σ^2 の $100(1-\alpha)$%信頼区間：　$\dfrac{(n-1)u^2}{\chi^2{}_{\frac{\alpha}{2}}(n-1)}<\sigma^2<\dfrac{(n-1)u^2}{\chi^2{}_{1-\frac{\alpha}{2}}(n-1)}$

比率 $\boldsymbol{p}$ の信頼区間（大標本）

p の $100(1-\alpha)$%信頼区間：　$\hat{p}-z_{\frac{\alpha}{2}}\sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}}<p<\hat{p}+z_{\frac{\alpha}{2}}\sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}}$

2 つの比率の差の信頼区間（大標本）

p_1-p_2 の $100(1-\alpha)$%信頼区間：

$$\hat{p}_1-\hat{p}_2-z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p}_1)}{n_1}+\frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}<p_1-p_2$$

$$<\hat{p}_1-\hat{p}_2+z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1}+\frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

例　　題

例題 1　（平均と分散の不偏推定値）

研究者は，ある母集団の平均と分散について推測をしたい．そのため母集団から無作為抽出によって，次の 10 個の標本を得た．

18, 25, 17, 20, 31, 28, 24, 23, 21, 24

このとき，（a）平均 μ,（b）分散 σ^2 の不偏推定値を求めよ．

解　まず，$\sum x_i = 18+25+\cdots+24=231$

$$\sum x_i{}^2 = 18^2+25^2+\cdots+24^2=5505$$

（a）　μ の不偏推定量は標本平均 $\overline{X}$ であるから，μ の不偏推定値は

$$\bar{x}=\frac{1}{n}\sum x_i=\frac{1}{10}\times 231=\mathbf{23.1}$$

（b）　σ^2 の不偏推定量は標本不偏分散 U^2 であるから，σ^2 の不偏推定値は

$$u^2=\frac{1}{n-1}\left\{\sum x_i{}^2-\frac{1}{n}\left(\sum x_i\right)^2\right\}$$

$$=\frac{1}{10-1}\left\{5505-\frac{1}{10}\times(231)^2\right\}\fallingdotseq\mathbf{18.77}$$

例題 2　**（より有効な推定量）**

平均 μ，分散 σ^2 の母集団からの大きさ 3 の無作為標本を X_1，X_2，X_3 とする．そのとき

（a）　$T_1=X_1+X_2-X_3$

（b）　$T_2=\dfrac{1}{4}(X_1+2X_2+X_3)$

（c）　$T_3=\dfrac{1}{3}(X_1+X_2+X_3)$

はいずれも μ の不偏推定量で，3 つの中では T_3 が最小の分散をもつことを示せ．

解　X_1，X_2，X_3 は無作為標本であるから，それらは互いに独立で，元の母集団と同じ平均 μ，同じ分散 σ^2 をもつ．

$$E(X_i)=\mu,\qquad V(X_i)=E(X_i-\mu)^2=\sigma^2\qquad(i=1,2,3)$$

したがって

（a）　$E(T_1)=E(X_1)+E(X_2)-E(X_3)=\mu+\mu-\mu=\mu$

$$V(T_1)=V(X_1)+V(X_2)+V(X_3)=\sigma^2+\sigma^2+\sigma^2=3\sigma^2$$

（b）　$E(T_2)=\dfrac{1}{4}\{E(X_1)+2E(X_2)+E(X_3)\}=\dfrac{1}{4}\times 4\mu=\mu$

$$V(T_2)=\frac{1}{4^2}\{V(X_1)+2^2V(X_2)+V(X_3)\}=\frac{1}{4^2}\times 6\sigma^2=\frac{3}{8}\sigma^2$$

（c）　$E(T_3)=\dfrac{1}{3}\{E(X_1)+E(X_2)+E(X_3)\}=\mu$

$$V(T_3)=\frac{1}{3^2}\times 3\sigma^2=\frac{1}{3}\sigma^2$$

これより

$$V(T_3) < V(T_2) < V(T_1)$$

ゆえに，T_3 は最小の分散をもつ．

例題 3　**（分散の不偏推定量）**

平均 μ，分散 σ^2 の母集団からの大きさ n の標本を，$X_1, X_2, \cdots, X_n$ とするとき，

$$U^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

は，σ^2 の不偏推定量であることを示せ．

解　標本変量 $X_1, X_2, \cdots, X_n$ は互いに独立で，その分布は母集団の分布と同じであるから，

$$E[X_i] = \mu, \quad E(X_i - \mu)^2 = \sigma^2 \qquad (i = 1, 2, \cdots, n)$$

$$\begin{aligned}
E(U^2) &= E\left[\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2\right] \\
&= \frac{1}{n-1}E\left[\sum(X_i{}^2 - 2X_i\overline{X} + \overline{X}^2)\right] \\
&= \frac{1}{n-1}E(\sum X_i{}^2 - 2\overline{X}\sum X_i + \sum \overline{X}^2) \\
&= \frac{1}{n-1}E(\sum X_i{}^2 - 2n\overline{X}^2 + n\overline{X}^2) \qquad \left(\overline{X} = \frac{1}{n}\sum X_i \text{ より}\right) \\
&= \frac{1}{n-1}E(\sum X_i{}^2 - n\overline{X}^2) \\
&= \frac{1}{n-1}\left[\sum_{i=1}^{n} E(X_i{}^2) - nE(\overline{X}^2)\right]
\end{aligned}$$

ここで，

$$\sigma^2 = E(X^2) - \mu^2 \qquad \Rightarrow \quad E(X^2) = \sigma^2 + \mu^2$$

$$\frac{\sigma^2}{n} = V(\overline{X}) = E(\overline{X}^2) - \mu^2 \qquad \Rightarrow \quad E(\overline{X}^2) = \frac{\sigma^2}{n} + \mu^2$$

これらを上の式に代入すれば

$$E(U^2) = \frac{1}{n-1}\left[n(\sigma^2 + \mu^2) - n\left(\frac{\sigma^2}{n} + \mu^2\right)\right] = \sigma^2$$

ゆえに，U^2 は σ^2 の不偏推定量である．

例題 4 （有効推定量）

平均 μ，分散 σ^2 の母集団からとった 2 個の無作為標本を X_1，X_2 とするとき，aX_1+bX_2 が μ の不偏推定量で，かつ有効推定量となるような a，b の値を求めよ．

解 X_1，X_2 は無作為標本だから

$$E(X_1)=E(X_2)=\mu, \qquad V(X_1)=V(X_2)=\sigma^2$$

aX_1+bX_2 の期待値は

$$E(aX_1+bX_2)=aE(X_1)+bE(X_2)=(a+b)\mu$$

aX_1+bX_2 が不偏推定量であるためには，

$$E(aX_1+bX_2)=\mu \qquad \therefore \quad a+b=1$$

$$V(aX_1+bX_2)=a^2V(X_1)+b^2V(X_2) \qquad (X_1 \text{ と } X_2 \text{ は独立だから})$$
$$=(a^2+b^2)\sigma^2$$

$V(aX_1+bX_2)$ を最小にする a，b の値は

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

より，$a=\boldsymbol{\frac{1}{2}}$，$b=1-a=\boldsymbol{\frac{1}{2}}$.

よって，標本平均 $\dfrac{X_1+X_2}{2}$ は μ の不偏で，かつ有効な推定量である．

例題 5 （標本数 n の決定）

ある特定銘柄の洗剤を使っている消費者の比率を推定したい．次の場合，少なくとも何個の標本が必要か．

（a） 母集団比率 p は 0.8 と 0.9 の範囲にあることがわかっていて，少なくとも 0.99 の確率で，標本比率 $\hat{p}$ と母集団比率 p との差を 0.02 以下にしたい場合．

（b） 母集団比率 p について何もわかっていなくて，少なくとも 0.95 の確率で標本比率 $\hat{p}$ と母集団比率 p との差を 0.03 以下にしたい場合．

解 （a） 標本数 n は大きくて，近似的に

$$Z=\frac{\hat{p}-p}{\sqrt{\frac{pq}{n}}}\sim N(0.1)\qquad (q=1-p)$$

が成り立つとすれば,

$$P(|Z|<2.576)=0.99$$

$$P\left(|\hat{p}-p|<2.576\sqrt{\frac{pq}{n}}\right)=0.99$$

題意より,

$$P(|\hat{p}-p|<0.02)\geqq 0.99$$

であるから，求める n は

$$2.576\sqrt{\frac{pq}{n}}\leqq 0.02$$

$$\therefore\quad n\geqq\left(\frac{2.576}{0.02}\right)^2 pq$$

$pq=p(1-p)$ は，$0.8\leqq p\leqq 0.9$ では $p=0.8$ のとき最大となるから

$$n\geqq\left(\frac{2.576}{0.02}\right)^2\times 0.8\times 0.2\fallingdotseq 2654.3$$

よって，**2655 個以上**とればよい．

（b） (a) と同様に解けば,

$$n\geqq\left(\frac{1.96}{0.03}\right)^2 pq$$

$pq=p(1-p)=\frac{1}{4}-\left(\frac{1}{2}-p\right)^2$ は $0\leqq p\leqq 1$ では，$p=\frac{1}{2}$ のとき最大となるから

$$n\geqq\left(\frac{1.96}{0.03}\right)^2\times\frac{1}{2}\times\frac{1}{2}\fallingdotseq 1067.1$$

よって，**1068 個以上**とればよい．

例題 6 （最尤推定量）

密度関数が

$$f(x)=\begin{cases}\frac{1}{\theta^2}xe^{-\frac{x}{\theta}} & (x\geqq 0\,;\,\theta>0)\\ 0 & (\text{その他})\end{cases}$$

で与えられる母集団からの大きさ n の無作為標本を $X_1, X_2, \cdots, X_n$ とするとき,

（a） θ の最尤推定量 $\hat{\theta}$ を求めよ．

（b） (a) で求めた $\hat{\theta}$ は θ の不偏推定量であることを示せ．

解　（a）尤度関数 $L(\theta)$ は

$$L(\theta)=\prod_{i=1}^{n}\frac{1}{\theta^2}X_ie^{-\frac{X_i}{\theta}}=\theta^{-2n}(X_1X_2\cdots X_n)e^{-\frac{1}{\theta}\sum_{i=1}^{n}X_i}$$

両辺の対数をとれば

$$\log L(\theta)=-2n\log\theta+\sum_{i=1}^{n}\log X_i-\frac{1}{\theta}\sum_{i=1}^{n}X_i$$

よって，

$$\frac{d\log L(\theta)}{d\theta}=-\frac{2n}{\theta}+\frac{1}{\theta^2}\sum_{i=1}^{n}X_i=0$$

$$\therefore\quad \hat{\theta}=\boldsymbol{\frac{1}{2}\overline{X}}$$

（b）$E(\hat{\theta})=\dfrac{1}{2}E(\overline{X})=\dfrac{1}{2}E(X)$　　（6 章の定理 5 より）

$$E(X)=\frac{1}{\theta^2}\int_0^\infty x^2e^{-\frac{x}{\theta}}dx$$

$$=\frac{1}{\theta^2}\left\{\left[-x^2\theta e^{-\frac{x}{\theta}}\right]_0^\infty+2\theta\int_0^\infty xe^{-\frac{x}{\theta}}dx\right\}\quad（部分積分より）$$

$$=\frac{1}{\theta^2}(0+2\theta\cdot\theta^2)\quad\left(\because\ \int_0^\infty\frac{xe^{-\frac{x}{\theta}}}{\theta^2}dx=1\right)$$

$$=2\theta$$

よって，

$$E(\hat{\theta})=\theta$$

ゆえに，θ の最尤推定量 $\dfrac{1}{2}\overline{X}$ は θ の不偏推定量でもある．

例題 7　（平均の信頼区間（大標本））

ある学校の生徒 40 人を無作為に選び，1 週間にテレビを何時間視るかを聞いた．40 人の平均は 18.2 時間，標準偏差 5.4 時間であった．この学校の生徒のテレビ平均視聴時間 μ に対する

（a）95％信頼区間，

（b）98％信頼区間

を求めよ．

解　（a）$n=40$ は大標本であるから，μ の 95％信頼区間は

$$\bar{x}-1.96\frac{u}{\sqrt{n}}<\mu<\bar{x}+1.96\frac{u}{\sqrt{n}}$$

に

$$n=40,\quad \bar{x}=18.2,\quad u=\sqrt{\frac{n}{n-1}}\,s=\sqrt{\frac{40}{39}}\times 5.4=5.5$$

を代入して

$$18.2-1.96\times\frac{5.5}{\sqrt{40}}<\mu<18.2+1.96\times\frac{5.5}{\sqrt{40}}$$

$$\mathbf{16.5<\mu<19.9}$$

（b）　98%の場合も同様にして

$$\bar{x}-2.33\frac{u}{\sqrt{n}}<\mu<\bar{x}+2.33\frac{u}{\sqrt{n}}$$

$$18.2-2.33\times\frac{5.5}{\sqrt{40}}<\mu<18.2+2.33\times\frac{5.5}{\sqrt{40}}$$

$$\mathbf{16.2<\mu<20.2}$$

例題　8　**（平均の信頼区間（小標本））**

ベアリングの製造機械がある．この機械が作る9個のベアリングの直径（mm）を測定して，

7.01, 7.03, 6.96, 6.91, 6.96, 7.06, 7.02, 6.94, 6.93

を得た．直径は正規分布に従うと仮定して，ベアリングの平均直径 μ に対する

（a）　90%信頼区間,

（b）　99%信頼区間

を求めよ．

解　データから和と2乗和を求めると

$$\sum x_i=62.82,\qquad \sum x_i^2=438.5$$

ゆえに

$$\bar{x}=\frac{1}{n}\sum x_i=6.98$$

$$u^2=\frac{1}{n-1}\left\{\sum x_i^2-\frac{1}{n}\left(\sum x_i\right)^2\right\}=0.00265$$

$$\therefore\quad u=\sqrt{0.00265}\fallingdotseq 0.05$$

（a）　μ の90%の信頼区間は,

$$\bar{x}-t_{0.05}(8)\frac{u}{\sqrt{n}}<\mu<\bar{x}+t_{0.05}(8)\frac{u}{\sqrt{n}}$$

$n=9$, $\bar{x}=6.98$, $u=0.05$, $t_{0.05}(8)=1.86$ を代入すれば

$$6.98-1.86\times\frac{0.05}{\sqrt{9}}<\mu<6.98+1.86\times\frac{0.05}{\sqrt{9}}$$

$$\mathbf{6.95<\mu<7.01}$$

（b） 同様にして，$t_{0.005}(8)=3.355$ より，μ の 99%の信頼区間は

$$6.98-3.355\times\frac{0.05}{\sqrt{9}}<\mu<6.98+3.355\times\frac{0.05}{\sqrt{9}}$$

$$\mathbf{6.92<\mu<7.04}$$

例題 9 （2つの平均の差の信頼区間(大標本)）

銘柄Aの電球80個の寿命（時間）のデータから，平均1070，分散472が得られた．一方，銘柄Bの60個の電球の寿命のデータから，平均1042，分散366を得た．2つの銘柄の電球の平均寿命時間の差の90%信頼区間を求めよ．

解 電球A，Bの真の寿命をそれぞれ μ_A，μ_B とする．

与えられた情報から

$$n_A=80,\quad \bar{x}_A=1070,\quad s_A{}^2=472$$

$$n_B=60,\quad \bar{x}_B=1042,\quad s_B{}^2=366$$

$1-\alpha=0.90$，$\alpha=0.10$ より，

$$z_{\frac{\alpha}{2}}=z_{0.05}=1.645$$

n_A，n_B は十分大きいから，

$$u_A{}^2\fallingdotseq s_A{}^2,\quad u_B{}^2\fallingdotseq s_B{}^2$$

よって，大標本による平均の差の信頼区間の公式より

$$1070-1042-1.645\sqrt{\frac{472}{80}+\frac{366}{60}}<\mu_A-\mu_B<1072-1042+1.645\sqrt{\frac{472}{80}+\frac{366}{60}}$$

$$\mathbf{22.30<\mu_A-\mu_B<33.70}$$

例題 10 （分散の信頼区間）

平均0，分散 σ^2 の正規母集団から大きさ10の標本を抽出して，次の結果を得た．母分散 σ^2 の95%信頼区間を求めよ．

0.067	2.066	3.192	0.515	2.194
−0.727	−0.547	−3.537	−1.613	1.582

解　公式より，σ^2 の $100(1-\alpha)\%$信頼区間は

$$\frac{(n-1)u^2}{\chi^2{}_{\frac{\alpha}{2}}(n-1)}<\sigma^2<\frac{(n-1)u^2}{\chi^2{}_{1-\frac{\alpha}{2}}(n-1)}$$

与えられたデータから

$$\sum_{i=1}^{10}x_1=3.192,\qquad \sum_{i=1}^{10}x_i{}^2=37.983$$

よって，

$$9u^2=\sum x_i{}^2-\frac{(\sum x_i)^2}{10}$$

$$=37.983-\frac{3.192^2}{10}=36.964$$

χ^2 分布表より，

$$\chi^2_{0.025}(9)=19.02,\qquad \chi^2_{0.975}(9)=2.70$$

であるから，σ^2 の 95%信頼区間は

$$\frac{36.964}{19.02}<\sigma^2<\frac{36.964}{2.70}$$

$$\therefore\quad \mathbf{1.94}<\sigma^2<\mathbf{13.69}$$

例題 11　（比率の信頼区間）

300 人の喫煙者に，ふだんすっているタバコの銘柄を質問したところ，72 人が銘柄 X を答えた．全喫煙者のうち，銘柄 X の常用者の比率の 95% 信頼区間を求めよ．

解　銘柄 X の常用者の比率を p とする．

題意より，$n=300$，$x=72$ だから，

$$\hat{p}=\frac{x}{n}=\frac{72}{300}=0.24$$

$z_{0.025}=1.96$ だから，p の 95%信頼区間は，

$$0.24-1.96\sqrt{\frac{0.24\times0.76}{300}}<p<0.24+1.96\sqrt{\frac{0.24\times0.76}{300}}$$

$$\mathbf{0.19}<p<\mathbf{0.29}$$

例題 12 (2つの比率の差の信頼区間)

ハウス内での園芸実験で，品種Aの種子100個と，品種Bの種子150個をまき，同一条件下で実験を行った．3週間後，品種Aは85個が発芽し，品種Bは114個が発芽した．両品種の発芽率の差に対する95%信頼区間を求めよ．

解　品種A，Bの母集団発芽率をそれぞれ p_1，p_2 とする．
与えられた情報から，

$$n_1=100,\quad x_1=85,\quad \hat{p}_1=\frac{85}{100}=0.85$$

$$n_2=150,\quad x_2=114,\quad \hat{p}_2=\frac{114}{150}=0.76$$

$$\alpha=0.05,\quad z_{\frac{\alpha}{2}}=z_{0.025}=1.645$$

よって，p_1-p_2 の95%信頼区間は，

$$0.85-0.76-1.96\sqrt{\frac{0.85\times0.15}{100}+\frac{0.76\times0.24}{150}}<p_1-p_2$$

$$<0.85-0.76+1.96\sqrt{\frac{0.85\times0.15}{100}+\frac{0.76\times0.24}{150}}$$

$$\mathbf{-0.008}<p_1-p_2<\mathbf{0.188}$$

7章の問題

7.1　ある有名予備校主催の模擬試験を受けた某中学校の生徒200人の得点は次のようであった．

階級値	4.5	14.5	24.5	34.5	44.5	55.5	64.5	74.5	84.5	94.5
度　数	3	11	18	24	41	43	30	15	10	5

(a)　この試験を受けた全受験生の得点の平均と分散を推定せよ．

(b)　平均に対する95%信頼限界を求めよ．

(c)　$P(|\bar{x}-\mu|<2)=0.95$ を満たす最小の n の値を求めよ．

7.2　平均 μ，分散1の母集団からの大きさ n の無作為標本によって，μ^2 を推定したい．標本平均 $\overline{X}$ の2乗は，μ^2 の不偏推定量にならないことを示し，μ^2 の不偏推定量を導け．

7.3　確率変数 X の密度関数は

$$f(x)=2x/\theta^2 \qquad (0<x<\theta)$$

で，いま X について5個の観測値

$$0.6,\ 1.5,\ 0.8,\ 1.1,\ 1.3$$

が得られた．このとき，

（a）θ,　（b）θ^2

の不偏推定値を求めよ．

7.4　確率密度関数

$$f(x\,;\theta)=(1+\theta)x^{\theta} \qquad (0<x<1\,;\theta>-1)$$

を分布にもつ母集団からの，大きさ n の標本にもとずく θ の最尤推定量を求めよ．

7.5　新車100台について，ガソリン1 ℓ 当たり走行キロ数を，コースで測定した．この結果から，1 ℓ 当たり平均走行キロ数の95%信頼限界を求めよ．

走行キロ数	台数
17.50〜17.99	2
18.00〜18.49	5
18.50〜18.99	8
19.00〜19.49	13
19.50〜19.99	19
20.00〜20.49	25
20.50〜20.99	15
21.00〜21.49	7
21.50〜21.99	5
22.00〜22.49	0
22.50〜22.99	1
計	100

7.6　平均 μ，分散 σ^2 の正規母集団からの大きさ12の無作為標本より，

$$\sum_{i=1}^{12} x_i=72, \qquad \sum_{i=1}^{12} x_i{}^2=1620$$

を得た．

（a）μ と σ^2 の不偏推定値を求めよ．

（b）μ と σ^2 の95%信頼区間を求めよ．

7.7　天秤で，ある物の重さを繰返し測定するときの測定値は，その物の真の重さに等しい平均と標準偏差 0.5 mg の正規分布に従う．

A の 10 回の測定値の平均は 8.6 mg で，B の 15 回の測定値の平均は 6.8 mg であった．測定値は正規分布に従うと仮定して，A と B の真の重さの差に対する 95%信頼区間を求めよ．

7.8　ある機械は，粉末製品を x_1 g ずつとって容器に詰める．過去の経験から，x_1 は平均 $\mu_1=30$ g，標準偏差 $\sigma_1=0.4$ g の正規分布に従う．容器の重さ x_2 は，平均 $\mu_2=5.0$ g，標準偏差 $\sigma_2=0.3$ g の正規分布に従う．容器こみの製品 1 個当りの重さ x に対する 95%信頼区間を求めよ．また，この製品 10 個を 1 箱に詰めるとき，空箱 1 個の重さ x_3 は $\mu_3=35.0$ g，$\sigma_3=1.0$ g の正規分布に従うとして，1 箱の重さ y に対する 95%信頼区間を求めよ．

7.9　生徒が 1800 人の中学校で，無作為に選んだ 40 人の生徒のうち 14 人がめがねをかけていた．

（a）　この学校で，めがねをかけている生徒の比率，

（b）　この学校で，めがねをかけている生徒の数

の 95%信頼限界を求めよ．

7.10　ある政策に対する世論調査で，男子有権者は 1000 人中 750 人が賛成と答え，女子有権者は 800 人中 520 人が賛成と答えた．母集団における男子有権者と女子有権者の賛成率の差に対する 95%信頼区間を求めよ．

7.11　螢光管の寿命時間は，指数分布

$$f(x\,;\theta)=\theta e^{-\theta x} \qquad (x \geqq 0)$$

に従う．5 個のデータ

5.83, 12.99, 16.28, 2.88, 1.83

が与えられたとき，θ の最尤推定値を求めよ．

8

検　定

8-1　仮説の検定

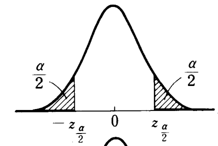
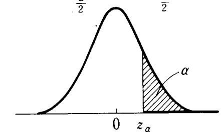
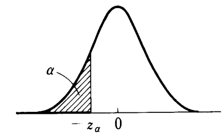
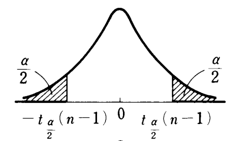
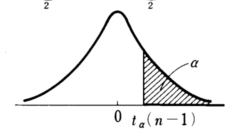
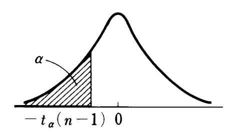
帰無仮説と対立仮説　母集団の母数に関するある主張または立言を**統計的仮説**または単に**仮説**という．母集団からとられた無作為標本の値によって母数に関する仮説の棄却か採択かを決めることを**仮説の検定**という．多くの場合，仮説は棄却されることを期待して設定する，たとえば，2 つの方法が優劣に関して差があると予想される場合には，「差がない」という仮説を設け，標本値によってこれを棄却しようとする．このように仮説はそれの棄却を期待して設定されるという意味で，**帰無仮説**とよばれる．帰無仮説との比較のために設けられる仮説で，帰無仮説が棄却されると受け入れられることになる仮説を**対立仮説**という．

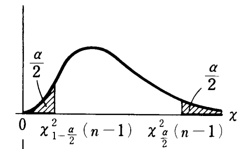
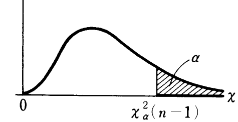
第 1 種の過誤と第 2 種の過誤　仮説の検定では 2 種類の誤りが考えられる．帰無仮説 H_0 が真のときこれを棄却する誤りを**第 1 種の過誤**といい，H_0 が偽のときこれを採択する誤りを**第 2 種の過誤**という．第 1 種の過誤を犯す確率を α，第 2 種の過誤を犯す確率を β で表す．α を**有意水準**，または**危険率**という．α の値としては，0.05，0.01 などがよく用いられる．

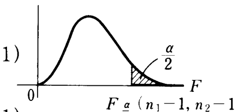
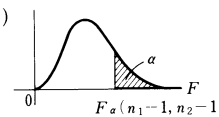
検定統計量と棄却域　実際の検定は，与えられた仮説に対して適当な統計量 T を選び，H_0 が真のときの T の標本分布を用いて行う．この T を**検定統計量**といい，H_0 が棄却されることになる T の実現値の範囲を**棄却域**という．

両側検定と片側検定　検定統計量 T の分布の左右の両裾を棄却域に選ぶ検定を**両側検定**．右側または左側の片裾を棄却域に選ぶ検定を**片側検定**といい，右裾のときを**右片側検定**，左裾のときを**左片側検定**という．ある仮説の検

標準的な検定の一覧表

仮　説	条　件	検定統計量	検定統計量の分布	棄却域
1. **平均の検定** (i) $H_0: \mu=\mu_0$ $H_1: \mu\neq\mu_0$	母集団の分布は正規分布 σ^2 は既知	$z=\dfrac{\bar{x}-\mu_0}{\dfrac{\sigma}{\sqrt{n}}}$ 大標本($n\geqq 30$)のときは左記の2条件は必要でなく，上の z の代りに $z=\dfrac{\bar{x}-\mu_0}{\dfrac{u}{\sqrt{n}}}$ を使う．ここで，u^2 は標本不偏分散である．n が大きいときは，$u^2 \fallingdotseq s^2$ だから，u の代りに s を用いてもよい．	標準正規分布	(i) $\lvert z\rvert > z_{\frac{\alpha}{2}}$ 
(ii) $H_0: \mu=\mu_0$ $H_1: \mu>\mu_0$				(ii) $z>z_\alpha$ 
(iii) $H_0: \mu=\mu_0$ $H_1: \mu<\mu_0$				(iii) $z<-z_\alpha$ 
2. **平均の検定** (小標本) (i) $H_0: \mu=\mu_0$ $H_1: \mu\neq\mu_0$	母集団の分布は正規分布 σ^2 は未知	$t=\dfrac{\bar{x}-\mu_0}{\dfrac{u}{\sqrt{n}}}$	自由度 $n-1$ の t 分布	(i) $\lvert t\rvert > t_{\frac{\alpha}{2}}(n-1)$ 
(ii) $H_0: \mu=\mu_0$ $H_1: \mu>\mu_0$				(ii) $t>t_\alpha(n-1)$ 
(iii) $H_0: \mu=\mu_0$ $H_1: \mu<\mu_0$				(iii) $t<-t_\alpha(n-1)$ 

3．**分散の検定** (i) $H_0: \sigma^2 = \sigma_0{}^2$ $H_1: \sigma^2 \neq \sigma_0{}^2$ (ii) $H_0: \sigma^2 = \sigma_0{}^2$ $H_1: \sigma^2 > \sigma_0{}^2$	母集団の分布は 正規分布	$\chi^2 = \dfrac{(n-1)u^2}{\sigma^2} = \dfrac{ns^2}{\sigma^2}$	自由度 $n-1$ の χ^2 分布	(i) $\chi^2 < \chi^2_{1-\frac{\alpha}{2}}(n-1)$ または $\chi^2 > \chi^2_{\frac{\alpha}{2}}(n-1)$ (ii) $\chi^2 > \chi^2_{\alpha}(n-1)$  
4．**比率の検定** (i) $H_0: p = p_0$ $H_1: p \neq p_0$ (ii) $H_0: p = p_0$ $H_1: p > p_0$ (iii) $H_0: p = p_0$ $H_1: p < p_0$	大標本 $n > 30$ $np > 5$, $nq > 5$	$z = \dfrac{\hat{p} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}}$ ここで，$\hat{p} = \dfrac{x}{n}$	標準正規分布	(i) $\lvert z \rvert > z_{\frac{\alpha}{2}}$ (ii) $z > z_\alpha$ (iii) $z < -z_\alpha$
5．**平均の差の検定** (i) $H_0: \mu_1 = \mu_2$ $H_1: \mu_1 \neq \mu_2$ (ii) $H_0: \mu_1 = \mu_2$ $H_1: \mu_1 > \mu_2$ (iii) $H_0: \mu_1 = \mu_2$ $H_1: \mu_1 < \mu_2$	母集団の分布は 正規分布 $\sigma_1{}^2$, $\sigma_2{}^2$ は既知	$z = \dfrac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{\sigma_1{}^2}{n_1} + \dfrac{\sigma_2{}^2}{n_2}}}$ 大標本(n_1, $n_2 \geqq 30$)のときは，左記の条件は必要でなく，上の z の代りに $z = \dfrac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{u_1{}^2}{n_1} + \dfrac{u_2{}^2}{n_2}}}$ を使う．n_1, n_2 が大きいときは，$u_1{}^2 \fallingdotseq s_1{}^2$, $u_2{}^2 \fallingdotseq s_2{}^2$ だから，$u_1{}^2$, $u_2{}^2$ の代りに $s_1{}^2$, $s_2{}^2$ を用いてもよい．	標準正規分布	(i) $\lvert z \rvert > z_{\frac{\alpha}{2}}$ (ii) $z > z_\alpha$ (iii) $z < -z_\alpha$

仮　説	条　件	検定統計量	検定統計量の分布	棄却域
6．平均の差の検定 （i）$H_0: \mu_1=\mu_2$ $H_1: \mu_1 \neq \mu_2$ （ii）$H_0: \mu_1=\mu_2$ $H_1: \mu_1>\mu_2$ （iii）$H_0: \mu_1=\mu_2$ $H_1: \mu_1<\mu_2$	母集団の分布は正規分布 $\sigma_1{}^2=\sigma_2{}^2=\sigma^2$ （等分散） で，σ^2 は未知	$t=\dfrac{\bar{x}_1-\bar{x}_2}{u\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}$ ここで， $u^2=\dfrac{(n_1-1)u_1{}^2+(n_2-1)u_2{}^2}{n_1+n_2-2}$	自由度 n_1+n_2-2 の t 分布	（i）$\|t\|>t_{\frac{\alpha}{2}}(n_1+n_2-2)$ （ii）$t>t_\alpha(n_1+n_2-2)$ （iii）$t<-t_\alpha(n_1+n_2-2)$
7．平均の差の検定 （対応のある場合） （i）$H_0: \mu_d=0$ $H_1: \mu_d \neq 0$ （ii）$H_0: \mu_d=0$ $H_1: \mu_d>0$ （iii）$H_0: \mu_d=0$ $H_1: \mu_d<0$	母集団の分布は正規分布	$t=\dfrac{\bar{d}\sqrt{n}}{u_d}$ ここで， $\bar{d}=\dfrac{1}{n}\sum d_i$ $u_d{}^2=\dfrac{1}{n-1}\sum(d_i-\bar{d})^2$	自由度 $n-1$ の t 分布	（i）$\|t\|>t_{\frac{\alpha}{2}}(n-1)$ （ii）$t>t_\alpha(n-1)$ （iii）$t<-t_\alpha(n-1)$
8．比率の差の検定 （i）$H_0: p_1=p_2$ $H_1: p_1 \neq p_2$ （ii）$H_0: p_1=p_2$ $H_1: p_1>p_2$ （iii）$H_0: p_1=p_2$ $H_1: p_1<p_2$	大標本 $n_1,\ n_2>30$ $n_1p_1>5,\ n_1q_1>5$ $n_2p_2>5,\ n_2q_2>5$	$z=\dfrac{\hat{p}_1-\hat{p}_2}{\sqrt{\hat{p}\hat{q}\left(\dfrac{1}{n_1}+\dfrac{1}{n_2}\right)}}$ ここで， $\hat{p}=\dfrac{x_1+x_2}{n_1+n_2}$	標準正規分布	（i）$\|z\|>z_{\frac{\alpha}{2}}$ （ii）$z>z_\alpha$ （ii）$z<-z_\alpha$
9．等分散の検定 （i）$H_0: \sigma_1{}^2=\sigma_2{}^2$ $H_1: \sigma_1{}^2 \neq \sigma_2{}^2$ （ii）$H_0: \sigma_1{}^2=\sigma_2{}^2$ $H_1: \sigma_1{}^2>\sigma_2{}^2$	母集団の分布は正規分布	$F=\dfrac{u_1{}^2,\ u_2{}^2\text{の大きい方}}{u_1{}^2,\ u_2{}^2\text{の小さい方}}$ $F=\dfrac{u_1{}^2}{u_2{}^2}$	自由度（$n_1-1,\ n_2-1$）の F 分布，または自由度（$n_2-1,\ n_2-1$）の F 分布 自由度（$n_1-1,\ n_2-1$）の F 分布	有意水準 α で， （i）$u_1{}^2>u_2{}^2$ ならば $F>F_{\frac{\alpha}{2}}(n_1-1,\ n_2-1)$ $u_2{}^2>u_1{}^2$ ならば $F>F_{\frac{\alpha}{2}}(n_2-1,\ n_1-1)$  （ii）$F>F_\alpha(n_1-1,\ n_2-1)$ 

定を両側検定で行うか，片側検定で行うかは，どんな対立仮説に関心をもつかによって決まる．たとえば，母集団比率 p の検定で，帰無仮説，$p=0.5$ に対して，対立仮説が $p\neq0.5$ ならば両側検定，$p>0.5$ ならば右側検定となる．

検定の一般的な手順

(i)　帰無仮説 H_0 と対立仮説 H_1 を定める．

(ii)　有意水準 α の値をきめる．

(iii)　検定統計量 T と棄却域 W を選ぶ．

(iv)　与えられたデータから T の実現値 T_0 を求めよ．

(v)　$T_0\in W$ ならば H_0 を棄却し，$T_0\notin W$ ならば H_0 を採択する．

例　題

例題 1 （平均値の検定）

正規母集団の平均値の検定において

(a)　$\bar{x}=38$, $\sigma=15$, $n=100$, $\alpha=0.05$ のとき，
$H_0: \mu=40$,　$H_1: \mu\neq40$　の検定をせよ．

(b)　$\bar{x}=240$, $\sigma=20$, $n=16$, $\alpha=0.05$ のとき，
$H_0: \mu=250$,　$H_1: \mu<250$　の検定をせよ．

(c)　$\bar{x}=51.52$, $s=2.13$, $n=14$, $\alpha=0.01$ のとき，
$H_0: \mu=50$,　$H_1: \mu>50$　の検定をせよ．

(d)　$\bar{x}=22.31$, $\bar{x}_2=21.54$, $s_1=3.8$, $s_2=3.2$, $n_1=50$, $n_2=40$, $\alpha=0.05$ のとき，
$H_0: \mu_1=\mu_2$,　$H_1: \mu_1\neq\mu_2$　の検定をせよ．

解　(a)　$$z=\frac{\bar{x}-\mu_0}{\frac{\sigma}{\sqrt{n}}}=\frac{38-40}{\frac{15}{\sqrt{100}}}=-1.33$$

両側検定だから，$z_{0.025}=1.96$ より棄却域は $|z|>1.96$.

$z=-1.33$ は棄却域に落ちないので，H_0 は**採択**.

(b)　$$z=\frac{\bar{x}-\mu_0}{\frac{\sigma}{\sqrt{n}}}=\frac{240-250}{\frac{20}{\sqrt{16}}}=-2$$

片側検定だから，$z_{0.05}=1.645$ より棄却域は $z<-1.645$

$z=-2$ は棄却域に落ちるから，H_0 は**棄却**.

（c）　$u=\sqrt{\frac{n}{n-1}s^2}=\sqrt{\frac{14}{14-1}\times 2.13^2}=2.29$

$$t=\frac{\bar{x}-\mu_0}{\frac{u}{\sqrt{n}}}=\frac{51.52-50}{\frac{2.29}{\sqrt{14}}}=2.48$$

$\alpha=0.01$, $\nu=14-1=13$. 片側検定だから，$t_{0.01}(13)=2.65$.
ゆえに棄却域は $t>2.65$. $t=2.48$ は棄却域に落ちないから H_0 は**採択**.

（d）　大標本の場合の 2 つの平均値の差の検定であるから

$$z=\frac{\bar{x}_1-\bar{x}_2}{\sqrt{\frac{s_1{}^2}{n_1}+\frac{s_2{}^2}{n_2}}}=\frac{22.31-21.54}{\sqrt{\frac{3.8^2}{50}+\frac{3.2^2}{40}}}=1.04$$

$\alpha=0.05$ の両側検定だから，棄却域は $|z|>1.96$.
$z=1.04$ は棄却域に落ちないから，H_0 は**採択**.

例題 2　**（平均値の検定（σ : 既知））**

ある機械が袋に詰める砂糖の重さは，$\mu=100$ g, $\sigma=5$ g の正規分布に従うように調整される．機械が正しく調整されているかどうかを確かめるために，9 個の袋の無作為標本をとって砂糖の重さを測ったら，平均 $\bar{x}$ は 102.4 g であった．この機械は正しく調整されているか，5%有意水準で検定せよ．

解　この機械が詰める袋の中の砂糖の重さを X とすると，$X\sim N(\mu, 5^2)$. 調整されているという帰無仮説を調整はうまくいっていないという対立仮説に対して行う検定であるから，両側検定である．よって

$$H_0: \mu=100$$
$$H_1: \mu\neq 100 \qquad \text{（両側検定）}$$

$n=9$, $\sigma=5$ で，$\bar{x}=102.4$ であるから，

$$z=\frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}}=\frac{102.4-100}{\frac{5}{\sqrt{9}}}=1.44$$

両側検定で，$\alpha=0.05$ であるから，$z_{\frac{\alpha}{2}}=z_{0.025}=1.96$.
よって，棄却域は

$$|z|>1.96$$

$z=1.44$ は棄却域に落らないから，仮説は採択される．
したがって，この機械は**正しく調整されている**と判断される．

例題 3　(平均値の検定 (小標本))

ある技師がニッケルの融点を 9 回測定して，次の値を得た (単位：°C).

1475	1420	1433	1452	1441
1466	1432	1453	1414	

この結果はニッケルの真の融点とされている 1455 °C に矛盾しないという仮説を有意水準 5% で検定せよ.

解　題意より，この問題は両側検定が適当である.

$$H_0 : \mu = 1455$$

$$H_1 : \mu \neq 1455$$

$n=9$, $\sum x_i = 12986$, $\sum x_i{}^2 = 18740684$ より

$$\bar{x} = \frac{12986}{9} = 1442.9, \qquad u^2 = \frac{1}{8}\left\{18740684 - \frac{12986^2}{9}\right\} = 416.1$$

$\alpha = 0.05$, $\nu = 9-1 = 8$. t 分布表より，$t_{0.025}(8) = 2.306$ であるから，棄却域は

$$|t| > 2.306$$

与えられた標本値から検定統計量の値は

$$t = \frac{\bar{x} - \mu_0}{\sqrt{\dfrac{u^2}{n}}} = \frac{1442.9 - 1455}{\sqrt{\dfrac{416.1}{9}}} = -1.78$$

この値は棄却域に落ちないから，H_0 は採択される. したがって，この実験結果は公表されているニッケルの融点の値に**矛盾していない**.

例題 4　(2 つの平均値の差の検定 (大標本))

ある年行われたレスリング選手権大会のある重量クラス 36 試合での勝者の平均体重は 64.5 kg，標準偏差は 3.2 kg で，敗者の平均体重は 62.8 kg，標準偏差は 2.5 kg であった. 選手の体重は試合の結果に影響するといえるか.「体重は試合の結果に影響しない」という帰無仮説を立て，選手の体重の分布は正規分布に従うと仮定して，5% 有意水準でこの仮説を検定せよ.

解　体重が試合の結果に影響しないということは，言い換えると「すべての勝者の平均体重 μ_1 とすべての敗者の平均体重 μ_2 との間に差がない」ということである. 体重の多い方が有利であると考えられるので，ここでは右側検定で検定を行うことにする.

$$H_0 : \mu_1 = \mu_2$$
$$H_1 : \mu_1 > \mu_2$$

$n_1 = n_2 = 36$ であるから，これは大標本による 2 つの平均値の差の検定である．

$\bar{x}_1 = 64.5$，$s_1 = 3.2$，$\bar{x}_2 = 62.8$，$s_2 = 2.5$ より

$$z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{s_1{}^2}{n_1} + \dfrac{s_2{}^2}{n_2}}} = \frac{64.5 - 62.8}{\sqrt{\dfrac{3.2^2}{36} + \dfrac{2.5^2}{36}}} = 2.51$$

$\alpha = 0.05$，片側検定であるから，$z_{0.05} = 1.645$，よって，棄却域は $z > 1.645$，$z = 2.51 > 1.645$ だから，仮説 H_0 は棄却される．

したがって，体重はレスリングの試合の結果に**影響する**．

例題 5　(**2 つの平均値の差の検定（小標本）**)

ある級は 20 人の男子生徒と 18 人の女子生徒からなる．男子生徒の数学のテストの得点は平均 65 点，標準偏差 15 点で，女子生徒の得点は平均 70 点，標準偏差 12 点であった．男女生徒の平均点の差は有意であるか．5% 有意水準で検定せよ．

解　男女の一方が特に成績がよいとは考えられないから，この問題は両側検定で行うのが適当である．よって，

$$H_0 : \mu_1 = \mu_2$$
$$H_1 : \mu_1 \neq \mu_2$$

ここで，μ_1 と μ_2 はそれぞれ男子と女子の得点の母平均である．

$\bar{x}_1 = 65$，$s_1 = 15$，$\bar{x}_2 = 70$，$s_2 = 12$，$n_1 = 20$，$n_2 = 18$ であるから

$$u^2 = \frac{(n_1 - 1)u_1{}^2 + (n_2 - 1)u_2{}^2}{n_1 + n_2 - 2}$$
$$= \frac{n_1 s_1{}^2 + n_2 s_2{}^2}{n_1 + n_2 - 2} = \frac{20 \times 15^2 + 18 \times 12^2}{20 + 18 - 2} = 197$$
$$\therefore \quad u \fallingdotseq 14.0$$

$$t = \frac{\bar{x}_1 - \bar{x}_2}{u\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}} = \frac{65 - 70}{14.0\sqrt{\dfrac{1}{20} + \dfrac{1}{18}}} \fallingdotseq -1.10$$

$\alpha = 0.05$，$\nu = 20 + 18 - 2 = 36$，$t_{0.025}(36) \fallingdotseq 2.03$ より，棄却域は

$$|t| > 2.03$$

$t = -1.10$ は棄却域に落ちていないから，H_0 は採択される．

したがって，男女生徒の成績の差は**有意とは認められない**．

例題 6 （平均値の差の検定（対応のある場合））

あるダイエット法が体重の減量に効果があるかどうか調べる実験に，10人の女性が参加した．この療法に入る前と，2ヵ月間試みた後の体重(kg)を測定して次の結果を得た．

女性	1	2	3	4	5	6	7	8	9	10
前	51.0	55.2	52.6	61.2	55.4	57.1	58.6	61.4	57.8	67.6
後	51.8	55.6	50.9	59.6	54.5	56.4	58.0	60.3	56.9	66.0

体重は正規分布に従うとして，このダイエット法は減量に効果があるかどうかを5%有意水準で検定せよ．

解　ダイエットに入る前と後の体重の差 d は

$$-0.8 \quad -0.4 \quad 1.7 \quad 1.6 \quad 0.9 \quad 0.7 \quad 0.6 \quad 1.1 \quad 0.9 \quad 1.6$$

$n=10$, $\sum d_i=7.9$, $\sum d_i{}^2=12.49$ より

$$\text{標本平均}\quad \bar{d}=\frac{7.9}{10}=0.79$$

$$\text{標本不偏分散}\quad u_d{}^2=\frac{1}{10-1}\left\{12.49-\frac{7.9^2}{10}\right\}=0.69$$

ここでの仮説は，題意より，

$$H_0 : \mu_d=0$$

$$H_1 : \mu_d>0$$

の右片側検定である．検定統計量の値は，

$$t=\frac{\bar{x}-\mu_0}{\sqrt{\dfrac{u^2}{n}}}=\frac{0.79-0}{\sqrt{\dfrac{0.69}{10}}}=3.01$$

$\alpha=0.05$, $\nu=10-1=9$ だから，t 分布表より $t_{0.05}(9)=1.833$．よって，棄却域は $t>1.833$．$t=3.01$ は棄却域に入るから，H_0 は棄却される．

したがって，この療法は減量の**効果があるといえる**．

例題 7 （分散の検定）

ガラスの仕切りから10個のガラスの標本をとり，ガラスの屈折率を測って次の値を得た．

1.77　1.79　1.78　1.79　1.79　1.76　1.80　1.76　1.79　1.80

ガラスの仕切りの受入れ検査では，屈折率の仕切り標準偏差が 0.008 を超えなければ，仕切りは合格で，超えれば不合格となる．このガラスの仕切りの合格，不合格を $\alpha=0.01$ で判定せよ．

解 データから

$$\sum x_i=1.77+1.79+\cdots+1.80=17.83$$

$$\sum x_i^2=1.77^2+1.79^2+\cdots+1.80^2=31.7929$$

$$s^2=\frac{1}{n}\left\{\sum x_i^2-\frac{(\sum x_i)^2}{n}\right\}=\frac{1}{10}\left\{31.7929-\frac{17.83^2}{10}\right\}=0.0002$$

仕切りでの屈折率の分散を σ^2 とすれば，題意より

$$H_0 : \sigma^2=0.008^2$$

$$H_1 : \sigma^2>0.008^2$$

$$\chi^2=\frac{ns^2}{\sigma^2}=\frac{10\times 0.0002}{0.008^2}=31.25$$

χ^2 分布表より，$\chi^2_{0.01}(9)=21.67$ であるから，この検定の棄却域は $\chi^2>21.67$．χ^2 の値 31.25 は棄却域に落ちるから，H_0 は棄却される．よって，仕切りは**不合格**と判定される．

例題 8（等分散の検定）

機械 I で作った錠剤を 10 個，機械 II で作った錠剤を 12 個とり，その重さ (g) を測って次の値を得た．

機械 I	0.374	0.380	0.365	0.399	0.388	0.383	0.311	0.398	0.363	0.373
機械 II	0.313	0.320	0.304	0.332	0.328	0.355	0.348	0.351	0.323	0.307
	0.342	0.374								

錠剤の重さは正規分布に従うと仮定して，これら 2 台の機械で作られる錠剤の重さの分散は等しいかどうかを有意水準 5% で検定せよ．

解 機械 I，II が作る錠剤の重さの母分散を σ_1^2, σ_2^2 で表せば，検定すべき仮説は，題意により

$$H_0 : \sigma_1^2=\sigma_2^2$$

$$H_1 : \sigma_1^2\neq\sigma_2^2$$

で，これは F 分布によって両側検定される．

データより

$$n_1=10,\quad \sum x_i=3.734,\quad \sum x_i{}^2=1.399958$$

$$n_2=12,\quad \sum y_i=3.997,\quad \sum y_i{}^2=1.336341$$

$$u_1{}^2=\frac{1}{n_1-1}\left\{\sum x_i{}^2-\frac{(\sum x_i)^2}{n}\right\}=\frac{1}{9}\left\{1.399958-\frac{3.734^2}{10}\right\}=0.000631$$

$$u_2{}^2=\frac{1}{n_2-1}\left\{\sum y_i{}^2-\frac{(\sum y_i)^2}{n}\right\}=\frac{1}{11}\left\{1.336341-\frac{3.997^2}{12}\right\}=0.000455$$

よって，

$$F=\frac{\max\{u_1{}^2,\ u_2{}^2\}}{\min\{u_1{}^2,\ u_2{}^2\}}=\frac{0.000631}{0.000455}=1.39$$

両側検定だから，$\alpha=0.05$ に対する自由度(9, 11)の F 分布の棄却限界値は，$F_{0.025}(9, 11)=3.60$（補間による）である．よって，$\alpha=0.05$ に対する両側検定の棄却域は

$$F>3.60$$

$F=1.39$ は棄却域に入らないから，H_0 は採択される．したがって，2 つの母分散は等しいとみなされる．

例題 9　**（比率の検定（大標本））**

サイコロを 300 回投げたら 6 の目が 72 回出た．このサイコロは偏っていると判断してよいか．$\alpha=0.05$ で検定せよ．

解　偏っていないとすれば，300 回のうち 6 の目の出る回数の平均は 50 回 $\left(=300\times\frac{1}{6}\right)$ のはずであるから，この場合の 72 回という結果は偶然にしては多過ぎないか，つまりこの結果は有意ではないのかというのが題意である．当然この場合は片側検定が使われる．このサイコロが 6 の目を出す確率を p とすると

$$H_0: p=\frac{1}{6}$$

$$H_1: p>\frac{1}{6}$$

$n=300$，$\hat{p}=\dfrac{x}{n}=\dfrac{72}{300}=0.24$．$\alpha=0.05$ で，片側検定だから，$z_{0.05}=1.645$．よって，棄却域は $z>1.645$．検定統計量の値は，

$$z=\frac{\hat{p}-p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}}=\frac{0.24-\dfrac{1}{6}}{\sqrt{\dfrac{\dfrac{1}{6}\times\dfrac{5}{6}}{300}}}=3.41$$

この値は棄却域に入るから，H_0 は棄却される．したがってこのサイコロは**偏っている**．

例題 10　(比率の検定（小標本))

スーパーマーケットで，6人の客に，手作りハムと工場製ハムを試食して味を比べてもらったら，6人中5人が手作りハムの方がおいしいと答えた．手作りハムは，工場製ハムより一般に好まれるといえるであろうか．消費者の両方のハムへの好みは同じであるという帰無仮説を5%有意水準で検定せよ．

解　消費者の母集団において，工場製より手作りのハムを好むという人の比率を p とする．題意より，この問題は比率 p の片側検定と考えられるから，仮説は

$$H_0 : p=\frac{1}{2}$$

$$H_1 : p>\frac{1}{2}$$

帰無仮説 H_0 が真のとき，6人中5人以上が手作りハムを好む確率は，2項分布 $B\left(6, \dfrac{1}{2}\right)$ より

$$P(X\geqq 5)=p(5)+p(6)=6\left(\frac{1}{2}\right)^5\left(\frac{1}{2}\right)+\left(\frac{1}{2}\right)^6=0.11$$

この確率は 0.05 より大きいから，5%有意水準で H_0 は棄却できない．

よって，消費者は工場製ハムより手作りハムをとくに好むとは**いえない**．

例題 11　(2つの比率の差の検定（大標本))

548人の男子高校生と582人の女子高校生に対し，どのような事柄について親と意見を異にするかの調査をした．以下の数値は，提示された項目

親と意見が一致しない項目	不一致の標本比率	
	男：$\hat{p}_1$	女：$\hat{p}_2$
友達の選択	0.243	0.272
大学への進学	0.375	0.298
洋服のスタイル	0.152	0.256
家事の手伝い	0.193	0.260
夜間の外出	0.458	0.486

に“イエス”と答えた男女生徒の標本比率である．ここにあげた項目のうち，どれが男女間で有意な差を示しているか．5%有意水準で，各項目ごとに比率の差の検定を行え．

解　各項目について，男子高校生と女子高校生の2つの母比率を比較する問題である．

項　目	$\lvert\hat{p}_1-\hat{p}_2\rvert$	$\lvert z\rvert=\dfrac{\lvert\hat{p}_1-\hat{p}_2\rvert}{\sqrt{\dfrac{\hat{p}_1(1-\hat{p}_1)}{n_1}+\dfrac{\hat{p}_2(1-\hat{p}_2)}{n_2}}}$
友達の選択	0.029	1.115
大学への進学	0.077	2.745 *
洋服のスタイル	0.104	4.385 *
家事の手伝い	0.067	2.702 *
夜間の外出	0.028	0.943

題意より，両側検定で行うのが適当である．

$z_{0.025}=1.96$ であるから，棄却域は，

$$\lvert z\rvert>1.96$$

z の値が棄却域に落ちて有意となるのは，上の表で＊印をつけた3つである．よって，男女間で有意な差のある項目は，

大学への進学，洋服のスタイル，家事の手伝い

である．

例題 12　**（第1種の過誤と第2種の過誤の確率）**

袋の中の球の構成は

（a）　赤球3個と黒球7個

または

（b）　赤球6個と黒球4個

のどちらかである．

(a) を帰無仮説 H_0, (b) を対立仮説 H_1 に選ぶ．この袋から2個の球を非復元抽出でとるとき，少なくとも1個が黒球であれば H_0 を採択し，それ以外のときは H_0 を棄却する．この検定の第1種と第2種の過誤の確率を求めよ．

解　とり出した2個の球のうち，黒球の数を X とする．
第1種の過誤は，H_0 が真のときこれを棄却する誤りであるから，その確率は

$$P(H_0\text{ を棄却}\,|\,H_0\text{ が真})=P(X=0\,|\,\text{赤球 3 個, 黒球 7 個})$$
$$=\frac{{}_3C_2\cdot{}_7C_0}{{}_{10}C_2}=\frac{3}{45}=\frac{1}{15}$$

一方，第2種の過誤は，H_1 が真のとき H_0 を採択する誤りであるから，その確率は

$$P(H_0\text{ を採択}\,|\,H_1\text{ が真})=P(X\neq 0\,|\,\text{赤球 6 個, 黒球 4 個})$$
$$=1-P(X=0\,|\,\text{赤球 6 個, 黒球 4 個})$$
$$=1-\frac{{}_6C_2\cdot{}_4C_0}{{}_{10}C_2}=\frac{2}{3}$$

8章の問題

8.1　比率の検定において

(a)　$n=120$，$x=90$，$\alpha=0.05$ のとき，
$H_0: p=0.7$，$H_1: p\neq 0.7$ の検定をせよ．

(b)　$n_1=420$，$x_1=118$，$n_2=530$，$x_2=121$，$\alpha=0.05$ のとき
$H_0: p_1=p_2$，$H_1: p_1>p_2$ の検定をせよ．

8.2　ある機械は重さ15gの鉛のおもりが作れるように設定されている．この機械が順調に作動しているかどうかを調べるため，12個のおもりの標本をとりその重さを測って，平均17g，標準偏差2.1gを得た．機械は順調に作動しているという仮説を5%有意水準で検定せよ．

8.3　ある機械は，長さ5.00cmの部品を切りとるように調整されている．この機械を使って7個の部品を切りとり，その長さを測定して次の数値を得た(単位：cm)．

4.98　5.01　4.96　4.95　4.94　5.02　4.96

この機械は所定の長さの部品を切りとっている，つまり機械は正しく調整されていると判断してよいか．正しく調整されているという仮説を，有意水準 $\alpha=0.05$ で検定せよ．

8.4　ある工場で生産された新型タイヤの中から，100個の無作為標本をとり，寿命試験を行った．その結果，タイヤの寿命は平均21000km，標準偏差

1500 km であった．これから，このタイヤの真の平均寿命は20000 kmを超えると主張してよいか，有意水準 $\alpha=0.05$ で検定せよ．

8.5 A，B 2台の機械により，粉乳を缶に詰める作業を行う．いま，両方の機械が詰めた170 g 入り缶をそれぞれ100個ずつとり，中味の量を計って次の結果を得た．

機械	標本の大きさ	標本平均（g）	標本標準偏差（g）
A	100	173.2	1.1
B	100	174.1	1.4

A，B 2台の機械が詰める粉乳の量の間に差があるといえるか．差がないという仮説を $\alpha=0.01$ で検定せよ．

8.6 2通りの牛の飼育法を比較する実験で，一卵生の子牛8対をとり，各対ごとにそれぞれ別の飼育法を割り当て一定期間飼育した．その後，これら牛の肉の味の良さを点数化して，次の結果を得た．

子牛の対	1	2	3	4	5	6	7	8
飼育法 A	57	61	61	68	59	69	71	57
飼育法 B	53	58	63	62	58	65	66	60

（a） 対応のある場合と考えて，2つの飼育法は味の良さについて異なるかどうか，を t 検定によって，5%有意水準で判定せよ．

（b） この検定に必要な仮定を述べよ．

8.7 ある母集団から20個のデータを抽出したら，標本分散が5.8であった．これから，もとの母集団の分散は10.0であるといえるだろうか．$\alpha=0.05$ で検定せよ．

8.8 2つの正規母集団からそれぞれ大きさ9と12の標本をとり，$u_1=7.5$，$u_2=4.2$ を得た．有意水準5%で次の仮説を検定せよ．

（a） $H_0: \sigma_1^2=\sigma_2^2$，　$H_1: \sigma_1^2\neq\sigma_2^2$

（b） $H_0: \sigma_1^2=\sigma_2^2$，　$H_1: \sigma_1^2>\sigma_2^2$

8.9　過去の経験によれば，ある高校から A 大学への合格率は 40% と考えられている．今年は 105 人受験して 36 人が合格した．今年の受験生は学力が劣っているといえるか．5% 有意水準で検定せよ．

8.10　ある都市のテレビ視聴率調査で，ある番組を男性は 200 人中 48 人が見たと答えたのに対し，女性は 300 人中 105 人が見たと答えた．この番組の視聴率は男女で差があるといえるだろうか．5% 有意水準で検定せよ．

8.11　1 枚の硬貨を 3 回投げるとき，おもての出る数を X とする．この硬貨がおもてを出す確率を p とし，X の値によって p に関する帰無仮説 $H_0: p=0.5$ を対立仮説 $H_1: p=0.8$ に対して検定するとき，検定の棄却域として，

（a）　$x=3$

または，

（b）　$x=0$

を選ぶ．そのとき，これら 2 つの検定に対する第 1 種の過誤の確率 α と第 2 種の過誤の確率 β をそれぞれ求めよ．

8.12　ある箱は

H_0 : 20 個の白球と 80 個の黒球を含む

H_1 : 50 個の白球と 50 個の黒球を含む

のいずれかである．仮説 H_0 を対立仮説 H_1 に対して検定するために，この箱から非復元抽出で 5 個の球を無作為にとり出す．このとき，5 個とも黒球が出たら，H_0 を採択し，それ以外ならば，H_1 を採択する．この検定の第 1 種の過誤と第 2 種の過誤の確率を求めよ．

8.13　つぼの中に多数の赤球と黒球が入っている．この中の赤球と黒球の割合は同じであるという仮説を検定するため，つぼから復元抽出で 25 個の球を無作為にとる．もし，その中の黒球の数が，15 個から 20 個の範囲内にあれば仮説を受け入れ，そうでないときは仮説を棄却する．仮説が実際正しいとき，この仮説を棄却する確率を求めよ．つぼの中の球の構成は，黒球が赤球より多いことが予想されているとする．このつぼから復元抽出で 64 個を無作為にとってその中の黒球の数を Y とするとき，Y を検定統計量とする有意水準 5% の棄却域を作れ．

9

カイ 2 乗検定

9-1　カイ 2 乗検定

一般に，帰無仮説の下である検定統計量が χ^2 分布に従う場合，この検定を**カイ 2 乗検定**という名でよぶことが多い．カイ 2 乗検定には，適合度検定や分割表での独立性の検定などがある．

9-2　適合度検定

母集団は互いに排反な k 個の級 $C_1, C_2, \cdots, C_k$ に分れているとし，この母集団から 1 個の標本をとるとき，それが $C_1, C_2, \cdots, C_k$ に入る確率を $p_1, p_2, \cdots, p_k$ $(\sum_{i=1}^{k} p_i=1)$ とする．いま，母集団から n 個の標本をとるとき，それらが $C_1, \cdots, C_k$ に入る**観測度数**を $n_1, n_2, \cdots, n_k$ $(\sum_{i=1}^{k} n_i=n)$ とすれば，これら級に入る**期待度数**は $m_1 \fallingdotseq np_1$, $m_2=np_2$, $\cdots$, $m_k=np_k$ となる．

級	C_1	C_2	$\cdots$	C_k	計
観測度数	n_1	n_2	$\cdots$	n_k	n
期待度数	m_1	m_2	$\cdots$	m_k	n

このとき，

帰無仮説 H_0：母集団の各級の確率(または確率分布)は

$$p_1=p_{10},\ p_2=p_{20},\ \cdots,\ p_k=p_{k0}$$

である

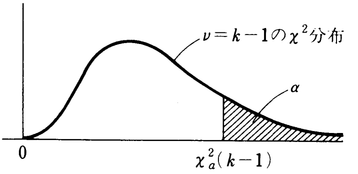
の検定に，検定統計量として

$$\chi^2=\sum_{i=1}^{k}\frac{(n_i-m_i)^2}{m_i}=\sum_{i=1}^{k}\frac{(n_i-n_ip_{i0})^2}{n_ip_{i0}}$$

が使われる．この統計量は n が大きく，各 m_i が5以上であれば H_0 の下で近似的に自由度 $\nu=k-1$ の χ^2 分布に従う．この検定を**適合度検定**という．有意水準が α %のこの検定の棄却域は

$$\chi^2>\chi^2_\alpha(k-1)$$

で与えられる．



期待度数の計算で，母集団の未知母数を推定することが必要な場合には，推定される母数が c 個ならば，このときのカイ2乗検定の棄却域は

$$\chi^2>\chi^2_\alpha(k-c-1)$$

となる．

9-3 独立性の検定

n 個の標本を2つの属性 A, B によって，次のような2元表に分割する．ここで，n_{ij} は級 (A_i, B_j) に入る標本の個数で，この表は **$r\times s$ 分割表**とよばれる．分類に用いる属性は定性的なものでも定量的なものでもよい．

$r\times s$ 分割表

A \ B	B_1	B_2	$\cdots$	B_s	計
A_1	n_{11}	n_{12}	$\cdots$	n_{1s}	$n_{1\cdot}$
A_2	n_{21}	n_{22}	$\cdots$	n_{2s}	$n_{2\cdot}$
$\vdots$	$\vdots$		$\ddots$	$\vdots$	$\vdots$
A_r	n_{r1}	n_{r2}	$\cdots$	n_{rs}	$n_{r\cdot}$
計	$n_{\cdot 1}$	$n_{\cdot 2}$	$\cdots$	$n_{\cdot s}$	n

$$n_{i\cdot}=\sum_{j=1}^{s}n_{ij},\qquad n_{\cdot j}=\sum_{i=1}^{r}n_{ij}$$

$$n=\sum_{i=1}^{r}\sum_{j=1}^{s}n_{ij}=\sum_{i=1}^{r}n_{i\cdot}=\sum_{j=1}^{s}n_{\cdot j}$$

1個の標本が級 (A_i, B_j) に入る確率を p_{ij}，級 A_i に入る確率を $p_{i\cdot}$，級 B_j に入る確率を $p_{\cdot j}$ とする．ここで検定すべき仮説は，2つの属性 A, B が独立であるという仮説，すなわち $P(A_i\cap B_j)=P(A_i)P(B_j)$ で，これは，

$$H_0: p_{ij}=p_{i\cdot}\,p_{\cdot j}\qquad (i=1, 2, \cdots, r\,; j=1, 2, \cdots, s)$$

で表される．H_0 が真の下で級 (A_i, B_j) に入る標本の期待度数 m_{ij} は，

$$m_{ij}=\frac{n_{i\cdot}\,n_{\cdot j}}{n}$$

で与えられる．よって，H_0 を検定する統計量は

$$\chi^2=\sum_{i=1}^{r}\sum_{j=1}^{s}\frac{\left(n_{ij}-\dfrac{n_{i\cdot}n_{\cdot j}}{n}\right)^2}{\dfrac{n_{i\cdot}n_{\cdot j}}{n}}$$

で，n が十分大きいならば，H_0 が真のとき，これは近似的に自由度 $\nu=(r-1)(s-1)$ の χ^2 分布に従う．よって，このカイ２乗検定の有意水準 α ％の棄却域は

$$\chi^2>\chi^2_\alpha((r-1)(s-1))$$

2×2 分割表　$r=s=2$ の場合を **2×2 分割表**という．このとき χ^2 は，次の式になる．

$$\chi^2=\frac{n(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)}$$

2×2 分割表

A \ B	B_1	B_2	計
A_1	a	b	$a+b$
A_2	c	d	$c+d$
計	$a+c$	$b+d$	n

イエーツの補正　2×2 の分割表で期待度数が小さい場合には χ^2 分布への近似をよくするために，$\dfrac{1}{2}$ だけずらして

$$\chi^2=\sum_{i=1}^{2}\sum_{j=1}^{2}\frac{\left(|n_{ij}-m_{ij}|-\dfrac{1}{2}\right)^2}{m_{ij}}$$

を使うのがよいとされている．これを**イエーツの補正**という．

例　題

例題 1　（一様分布の適合度検定）

観測者が測定器具の目盛りを読むとき，最後の桁の数字は目測で判読される．その際，特定の数字を好む傾向のあることが指適されている．いま，200 個の数字について，次の結果が得られた．この観測者は特別な数字を好む傾向があるだろうか．5％有意水準で検定せよ．

最後の桁の数字	0	1	2	3	4	5	6	7	8	9	計
観測度数	32	16	18	19	17	25	11	16	30	16	200

解　この場合の帰無仮説としては「観測者は特別な数字を好む傾向をもたない」をとるのが適当で，観測者が最後の桁の数字を i と読む確率を p_i とすると，この仮説は

$$H_0: p_0 = p_1 = \cdots = p_9 = \frac{1}{10}$$

で表される．H_0 が真であれば，期待度数は明らかに

$$m_i = np_{i0} = 200 \times \frac{1}{10} = 20 \qquad (i = 0, 1, \cdots, 9)$$

であるから，観測度数 n_i と期待度数 m_i を表にして示すと

	0	1	2	3	4	5	6	7	8	9	計
n_i	32	16	18	19	17	25	11	16	30	16	200
m_i	20	20	20	20	20	20	20	20	20	20	200

よって，

$$\chi^2 = \sum_i \frac{(n_i - m_i)^2}{m_i} = \frac{(32-20)^2}{20} + \frac{(16-20)^2}{20} + \cdots + \frac{(16-20)^2}{20} = 20.6$$

$\nu = 10 - 1 = 9$，$\alpha = 0.05$．χ^2 分布表より $\chi^2_{0.05}(9) = 16.92$．ゆえに，棄却域は

$$\chi^2 > 16.92$$

データから求めた χ^2 の値 20.6 は棄却域に入るから，H_0 は棄却される．

したがって，この観測者には特別な数字を好む傾向があると**いえる**．実際，この観測者は 0 と 8 を好むようである．

例題 2　**（適合度検定）**

人間の血液型は 4 種類で，その構成比率は $q^2 : p^2 + 2pq : r^2 + 2qr : 2pr$ であるという．ただし $p + q + r = 1$．いま，ある職業についている 770 人の血液型を調べて，

180, 360, 132, 98

なる観測度数を得た．これより，仮説「この職業人の血液型の分布は，$p = 0.4$，$q = 0.4$，$r = 0.2$ で定まる構成比率をもつ」を検定せよ．

解　仮説が正しいとして 4 種類の血液型の期待度数を求めれば

$$n_1 = nq^2 = 770 \times 0.4^2 = 123.2$$

$$n_2 = n(p^2 + 2pq) = 770 \times (0.4^2 + 2 \times 0.4 \times 0.4) = 369.6$$

$$n_3 = n(r^2+2qr) = 770\times(0.2^2+2\times0.4\times0.2) = 154.0$$
$$n_4 = n(2pr) = 770\times(2\times0.4\times0.2) = 123.2$$

観測度数と比較すると，

					計
n_i	180	360	132	98	770
m_i	123.2	369.6	154.0	123.2	770.0

$$\chi^2 = \frac{(180-123.2)^2}{123.2} + \frac{(360-369.6)^2}{369.6} + \frac{(132-154.0)^2}{154.0} + \frac{(98-123.2)^2}{123.2}$$
$$= 34.73$$

$\alpha=0.05$, $\nu=4-1=3$. χ^2 分布表より $\chi^2_{0.05}(3)=7.81$. よって棄却域は

$$\chi^2 > 7.81$$

χ^2 の値は棄却域に入っているから，仮説は棄却される．したがってこの職業人の血液型の分布は仮説が与える構成比率とは**異なる**ものである．

例題 3　**（ポアソン分布の適合度検定）**

大気中に浮遊するある微小な物質の量を推定するため，空間内にいくつかの点を選び，その点のまわりの単位体積内の粒子数を計測する．いま 300 点を選んで観測した結果，つぎのデータが得られた．

（a）　粒子の数の平均と分散を求めよ．

（b）　このデータにポアソン分布をあてはめ，その適合性を調べよ．

粒子の数	0	1	2	3	4	5	6 以上	計
観測度数	38	75	89	54	20	19	5	300

解　（a）　粒子数を x，度数を f とすると，平均は

$$\bar{x} = \frac{\sum x_i f_i}{n} = \frac{0\times38+1\times75+\cdots+6\times5}{300} = \frac{31}{15} \fallingdotseq \mathbf{2.07}$$

分散は

$$s^2 = \frac{\sum x_i^2 f_i}{n} - \bar{x}^2 = \frac{0^2\times38+1^2\times75+\cdots+6^2\times5}{300} - \left(\frac{31}{15}\right)^2 \fallingdotseq \mathbf{2.04}$$

平均と分散がほぼ等しいので，粒子の数の分布としてポアソン分布が予想される．

(b)　粒子の数は $\lambda=2.07$ のポアソン分布に従うという帰無仮説の下で，期待度数を次のように求める．

値 x	0	1	2	3	4	5	6 以上	計
確率 $e^{-2.07}\dfrac{2.07^x}{x!}$	0.13	0.26	0.27	0.19	0.10	0.04	0.01	1.0
期待度数＝300×確率	39	78	81	57	30	12	3	300
観測度数	38	75	89	54	20	19	5	300

これから

$$\chi^2=\frac{(38-40.6)^2}{40.6}+\frac{(75-81.2)^2}{81.2}+\cdots+\frac{(5-5.0)^2}{5.0}=9.39$$

$\alpha=0.05$，$\nu=7-1=6$ より $\chi^2_{0.05}(6)=12.59$．$\chi^2=9.39$ はこの値を超えないから，仮説は採択である．よってこのデータはポアソン分布に**適合している**．

例題 4　(2×2 分割表)

ある会社の社員 60 名に，タバコをすうかすわないかと，パチンコをするかしないかを調査した．その結果は次のようであった．タバコをすうこととパチンコをすることは独立かどうかを 5% 有意水準で検定せよ．

	パチンコをする	しない	
タバコをすう	9	3	12
すわない	18	30	48
	27	33	60

解　この問題の帰無仮説は，H_0：タバコとパチンコとは独立，である．H_0 の下での期待度数は

	パチンコ	しない
タバコ	$27\times12/60=5.4$	$33\times12/60=6.6$
すわない	$27\times48/60=21.6$	$33\times48/60=26.4$

イエーツの補正を施してカイ 2 乗検定を行う．

観測度数 n_i	期待度数 m_i	$n_i - m_i$	$\|n_i - m_i\| - 0.5$	$\dfrac{(\|n_i - m_i\| - 0.5)^2}{m_i}$
9	5.4	3.6	3.1	1,780
3	6.6	−3.6	3.1	1,456
18	21.6	−3.6	3.1	0.445
30	26.4	3.6	3.1	0.364
				計　4.045

$\alpha=0.05$, $\nu=1$ より $\chi^2_{0.05}(1)=3.84$ であるから，棄却域は $\chi^2>3.84$. よって帰無仮説は棄却される．したがってタバコとパチンコは**独立ではない**.

例題 5　**(2×4 分割表)**

アメリカでの調査によると，息子が父親の職業と同じ職業を選ぶかどうかを調べて，次の結果を得た．

息子の職業	父親の職業 医者	銀行員	教員	弁護士	計
父親と同じ職業	34	27	28	19	108
父親と異なる職業	166	123	152	81	522
計	200	150	180	100	630

「息子の職業選択と親の職業とは独立である」という仮説を 5%有意水準で検定せよ．

解　これは 2×4 分割表での独立性の検定の問題である．
各枡の期待度数を求めると，

$$m_{11}=\frac{200\times108}{630}=34.3,\qquad m_{12}=\frac{150\times108}{630}=25.7,$$

$$m_{13}=\frac{180\times108}{630}=30.9,\qquad m_{14}=\frac{100\times108}{630}=17.1,$$

$$m_{21}=\frac{200\times522}{630}=165.7,\qquad m_{22}=\frac{150\times522}{630}=124.3,$$

$$m_{23}=\frac{180\times522}{630}=149.1,\qquad m_{24}=\frac{100\times522}{630}=82.9$$

よって期待度数の表は

34.3	25.7	30.9	17.1
165.7	124.3	149.1	82.9

これから

$$\chi^2=\sum_{i,j}\frac{(n_{ij}-m_{ij})^2}{m_{ij}}=\frac{(34-34.3)^2}{34.3}+\frac{(27-25.7)^2}{25.7}+\cdots+\frac{(81-82.9)^2}{82.9}$$
$$=0.64$$

自由度 $\nu=(2-1)(4-1)=3$，$\chi^2_{0.05}(3)=7.81$ であるから，仮説は棄却されない．したがって息子の職業選択に親の職業は**関係しない**．

9章の問題

9.1　乱数サイを200回実際に振って，次の結果を得た．この乱数サイは正しいと認められるか．

目の数	0	1	2	3	4	5	6	7	8	9	計
度数	26	27	20	13	19	19	15	19	27	15	200

9.2　次の表は100個の乱数の標本であるといわれている．0から9までの各数字の度数はその期待度数と有意に異なるか．5%有意水準で検定せよ．

35230	66852	50395	59228	28896	48780	00845	39797	86339	57380
92264	95450	41210	66273	91350	52137	02829	62316	46155	16031

9.3　次の表は10騎馬兵団の20年間の記録で，馬にけられて死んだ兵士の数である．これはBortkewitchによって集められ，Fisherが引用した有名な例である．これに適当なポアソン分布をあてはめ，その適合性を論ぜよ．

死者の数	0	1	2	3	4	計
兵団数	109	65	22	3	1	200

9.4　流行性感冒の予防注射の効果を調べるため，流行性感冒にかかった人とかからなかった人について，それぞれ予防注射をしたか，しなかったかをきき，次の結果を得た．この予防注射は流行性感冒の予防に効果があるといえるか．1%有意水準で検定せよ．

注射	流行性感冒 かかった	かからなかった	計
した	32	44	76
しなかった	128	56	184
計	160	100	260

9.5 次の表はある教師による成績評価を A, B, C の 3 学部別に示したものである．これより，学部間に成績の優劣が認められるかどうかを有意水準 1%で検定せよ．

	A	B	C	計
優	8	12	6	26
良	25	32	10	67
可	10	9	14	33
不可	7	5	12	24
計	50	58	42	150

10

相関と回帰

10-1　相関係数

2つの変量 X, Y に関する n 対のデータ (x_1, y_1), (x_2, y_2), $\cdots$, (x_n, y_n) が与えられたとき，これら n 個の点を平面上にプロットした図を**散布図**という．散布図をみれば，X と Y の間に直線的関係があるか，あるとすれば，それはどの程度かなどがわかる．

相関係数　　X と Y の間の直線的関係の程度を表す統計的尺度を**相関係数**といい，

$$r=\frac{\frac{1}{n}\sum(x_i-\bar{x})(y_i-\bar{y})}{s_x s_y}$$

で定義する．ここで，$\bar{x}$, $\bar{y}$ は X, Y の平均，s_x, s_y は X, Y の標準偏差を表す．分子の $\frac{1}{n}\sum(x_i-\bar{x})(y_i-\bar{y})$ を X と Y の**共分散**という．

r の計算　　r の計算は，上の定義式より次の式が便利である．

$$r=\frac{n\sum x_iy_i-(\sum x_i)(\sum y_i)}{\sqrt{\{n\sum x_i^2-(\sum x_i)^2\}\{n\sum y_i^2-(\sum y_i)^2\}}}$$

r の性質　　(i)　$-1\leqq r\leqq 1$ が成り立つ．r の絶対値が1に近いときは，X, Y の間に強い直線関係があることを意味し，r の絶対値が0に近いときは，X, Y の間に直線的関係がほとんどないことを意味する．$r>0$ ならば，**正の相関**があるといい，$r<0$ ならば**負の相関**があるという．$|r|=1$ のとき**完全相関**，$r=0$ のとき**無相関**という．

(ii)　$\dfrac{x-a}{c}=u$, $\dfrac{y-b}{d}=v$ (a, b, c, d は任意定数) によって，(X, Y) を (U, V) に変換しても，相関係数の値は変らない．すなわち，

$$r_{xy}=r_{uv}$$

最小 2 乗法　散布図から X, Y 間に直線的関係のあることがわかったとき，これらのデータに直線をあてはめる問題が考えられる．これは直線を $y=a+bx$ とするとき，

$$\sum_{i=1}^{n}(y_i-a-bx_i)^2$$

を最小にするような a, b を求めるという考えで，直線を定める方法が使われる．このような方法を**最小 2 乗法**といい，この方法で求めた直線を**最小 2 乗直線**，または Y の X への**回帰直線**という．

順位相関係数　n 個の標本が 2 つの変量 X, Y についてそれぞれ順位がつけられているとき，これら 2 組の順位 $(x_1, x_2, \cdots, x_n)$ と $(y_1, y_2, \cdots, y_n)$ の間の相関係数は

$$r_S=1-\frac{6\sum d_i{}^2}{n(n^2-1)} \qquad (d_i=x_i-y_i)$$

となる．これを**スピアマンの順位相関係数**という．

10-2　相関係数の検定

母集団相関係数を ρ, 標本の大きさを n, 標本相関係数を r とする．

$\rho=0$ の検定　$H_0:\rho=0$

$H_1:\rho\neq0$　(両側検定)

の検定は，検定統計量

$$t=\sqrt{\frac{r^2(n-2)}{1-r^2}}$$

が，H_0 の下で自由度 $n-2$ の t 分布に従うことから，有意水準 α %の棄却域は

$$|t|>t_{\frac{\alpha}{2}}(n-2)$$

で与えられる．

$\rho=\rho_0$ の検定（大標本）　$H_0:\rho=\rho_0$

$H_1:\rho\neq\rho_0$　(両側検定)

の検定は，

$$z=\frac{1}{2}\log\frac{1+r}{1-r}, \qquad \mu=\frac{1}{2}\log\frac{1+\rho}{1-\rho}$$

とおくとき，検定統計量 z が H_0 の下で近似的に $N\left(\mu, \dfrac{1}{n-3}\right)$ に従うことから，有意水準 α %の棄却域は

$$\frac{|z-\mu|}{\frac{1}{\sqrt{n-3}}} \geqq z_{\frac{\alpha}{2}}$$

で与えられる．

回帰直線　　$y=a+bx$　または　$y-\bar{y}=b(x-\bar{x})$

ここで，$\bar{x}$，$\bar{y}$ は X，Y の平均で，$\bar{x}=\frac{1}{n}\sum x_i$，$\bar{y}=\frac{1}{n}\sum y_i$．また b，a は

$$b=\frac{\sum (x_i-\bar{x})(y_i-\bar{y})}{\sum (x_i-\bar{x})^2}=\frac{n\sum x_i y_i-(\sum x_i)(\sum y_i)}{n\sum x_i^2-(\sum x_i)^2}$$

$$a=\bar{y}-b\bar{x}$$

である．この直線を **Y の X への回帰直線**といい，a, b を**回帰係数**という．

10-3　線形回帰

2 変数 X，Y 間の構造に関して線形回帰モデルとよばれる確率モデルを仮定し，モデルの式に含まれている母数の推定や検定の問題を考える．

線形回帰モデル　　$x_1, x_2, \cdots, x_n$ を X の選ばれた値，$y_1, y_2, \cdots, y_n$ をこれら X の値に対する観測値とするとき，

$$y_i=\alpha+\beta x_i+\varepsilon_i \qquad (i=1, 2, \cdots, n)$$

と書ける．ここで，$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ は互いに独立で同一の正規分布 $N(0, \sigma^2)$ に従う誤差変量である．このモデルを**線形回帰モデル**という．このとき，X を**独立変数**，Y を**従属変数**という．

線形回帰モデルの下において，種々の推測問題が考えられる．

誤差分散 σ^2 の推定値　　σ^2 の不偏推定値は

$$s_e{}^2=\frac{1}{n-2}\sum (y_i-y_i')^2 \qquad (y_i'=\bar{y}+b(x_i-\bar{x}))$$

$$=\frac{1}{n-2}\{\sum (y_i-\bar{y})^2-b\sum (x_i-\bar{x})(y_i-\bar{y})\}$$

$$=\frac{1}{n-2}\{\sum y_i{}^2-a\sum y_i-b\sum x_i y_i\}$$

β の信頼限界　　母集団回帰係数 β の $100(1-\alpha)$％信頼限界は

$$b \pm t_{\frac{\alpha}{2}}(n-2)\frac{s_e}{\sqrt{\sum (x_i-\bar{x})^2}}$$

$\beta=0$ の検定　　$H_0: \beta=0$

$H_1: \beta \neq 0$　　（両側検定）

の検定は，検定統計量

$$t=\frac{b}{\dfrac{s_e}{\sqrt{\sum(x_i-\bar{x})^2}}}$$

が H_0 の下で自由度 $n-2$ の t 分布に従うことから，棄却域

$$|t|>t_{\frac{\alpha}{2}}(n-2)$$

で与えられる．

特定の x の値に対する Y の母平均の信頼限界　$X=x_0$ に対する Y の母平均の $100(1-\alpha)$%信頼限界は，

$$a+bx_0\pm t_{\frac{\alpha}{2}}(n-2)\,s_e\sqrt{\frac{1}{n}+\frac{(x_o-\bar{x})^2}{\sum(x_i-\bar{x})^2}}$$

で与えられる．

X の Y への回帰直線　この場合は，Y を独立変数，X を従属変数と考える．

$$x=a'+b'y \quad \text{または} \quad x-\bar{x}=b'(y-\bar{y})$$

ここで x，y は X，Y の平均，

$$b'=\frac{\sum(x_i-\bar{x})(y_i-\bar{y})}{\sum(y_i-\bar{y})^2}=\frac{n\sum x_iy_i-(\sum x_i)(\sum y_i)}{n\sum y_i^2-(\sum y_i)^2},\qquad a'=\bar{x}-b'\bar{y}$$

この直線を X の Y への**回帰直線**といい，a', b' を**回帰係数**という．

例　題

例題 1　**（相関係数）**

次に示す 10 人の学生の身長 (X) と体重 (Y) のデータから，相関係数を求めよ．

身長（x cm）	166	172	182	183	174	152	170	173	160	165
体重（y kg）	60.5	62.9	65.0	64.8	59.2	56.9	61.0	61.9	59.3	58.3

解　計算を簡略化するため，$u=x-170$，$v=y-60$ とおく．

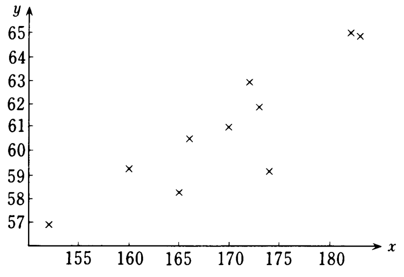
x	y	$u=x-170$	$v=y-60$	u^2	v^2	uv
166	60.5	-4	0.5	16	0.25	-2.0
172	62.9	2	2.9	4	8.41	5.8
182	65.0	12	5.0	144	25.00	60.0
183	64.8	13	4.8	169	23.04	62.4
174	59.2	4	-0.8	16	0.64	-3.2
152	56.9	-18	-3.1	324	9.61	55.8
170	61.0	0	1.0	0	1.00	0
173	61.9	3	1.9	9	3.61	5.7
160	59.3	-10	-0.7	100	0.49	7.0
165	58.3	-5	-1.7	25	2.89	8.5
計		-3	9.8	807	74.94	200

$n=10$, $\sum u_i=-3$, $\sum v_i=9.8$, $\sum u_i^2=807$, $\sum v_i^2=74.94$, $\sum u_iv_i=200$ を r の計算式に代入すれば

$$r_{xy}=r_{uv}=\frac{n\sum u_iv_i-(\sum u_i)(\sum v_i)}{\sqrt{\{n\sum u_i^2-(\sum u_i)^2\}\{n\sum v_i^2-(\sum v_i)^2\}}}$$

$$=\frac{10\times200-(-3)\times9.8}{\sqrt{(10\times807-(-3)^2)(10\times74.94-(9.8)^2)}}=\mathbf{0.88}$$

散布図を示すと



例題 2 (特別な散布図と相関係数)

次のデータに対して，相関係数 r を求めよ．それぞれについて，散布図をかき，得られた r の値について論評せよ．

(a)

X	1	2	3	4	5
Y	2	4	6	8	10

(b)

X	1	1	−1	−1
Y	1	−1	1	−1

(c)

X	−3	−2	−1	0	1	2	3
Y	8	3	0	−1	0	3	8

解

(a)　$n=5,\ \sum x_i=15,\ \sum y_i=30,\ \sum x_i^2=55,\ \sum y_i^2=220,\ \sum x_iy_i=110.$

$$r=\frac{n\sum x_iy_i-(\sum x_i)(\sum y_i)}{\sqrt{\{n\sum x_i^2-(\sum x_i)^2\}\{n\sum y_i^2-(\sum y_i^2)\}}}$$

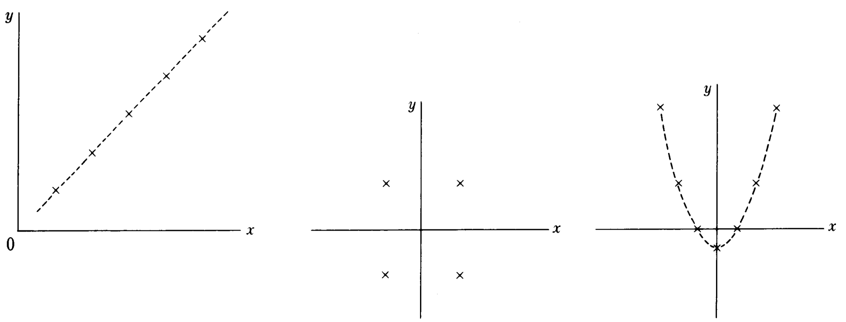
$$=\frac{5\times110-15\times30}{\sqrt{(5\times55-15^2)(5\times220-30^2)}}=\mathbf{1}$$

(b)　$n=4,\ \sum x_i=0,\ \sum y_i=0,\ \sum x_i^2=4,\ \sum y_i^2=4,\ \sum x_iy_i=0.$

$$r=\frac{4\times0-0\times0}{\sqrt{(4\times4-0^2)(4\times4-0^2)}}=\mathbf{0}$$

(c)　$n=7,\ \sum x_i=0,\ \sum y_i=-1,\ \sum x_i^2=28,\ \sum y_i^2=147,\ \sum x_iy_i=0.$

$$r=\frac{7\times0-0\times(-1)}{\sqrt{(7\times28-0^2)(7\times147-(-1)^2)}}=\mathbf{0}$$



(a)　　(b)　　(c)

(a)は各観測点が直線 $y=2x$ の上にのっているので，$r=+1$ になるのは当然.

(b)は，4 つの点が原点に関して対称な位置にあって，直線的関係は全くないので，$r=0$ となっている.

(c)は，7 個の点が放物線 $y=x^2-1$ の上にのっており，直線関係以外の放物線的関係なので，直線的傾向しか測り得ない相関係数は 0 となる.

例題 3 **（最小 2 乗法）**

ある計算アルゴリズムは，変数の個数 x と計算時間 y との間に

$$y=a+b\log_{10}x$$

の関係があるという．実際のデータは

個数 x	10	50	100	500	1000
時間 y	7.2	8.5	9.1	10.3	11.0

このとき，a と b の値を最小 2 乗法で求めよ．また散布図の上に求めた直線をかき入れよ．

解　いま $X=\log_{10}x$，$Y=y$ とおき，

$$Y=a+bX$$

の形にして，最小 2 乗法で，a，b の値を求める．

	X	Y	XY	X^2
	1	7.2	7.2	1
	1.7	8.5	14.45	2.89
	2	9.1	18.2	4
	2.7	10.3	27.81	7.29
	3	11.0	33.0	9
計	10.4	46.1	100.66	24.18

$$\overline{X}=\frac{10.4}{5}=2.08,\qquad \overline{Y}=\frac{46.1}{5}=9.22$$

$$b=\frac{n\sum X_iY_i-(\sum X_i)(\sum Y_i)}{n\sum X_i{}^2-(\sum X_i)^2}=\frac{5\times100.66-10.4\times46.1}{5\times24.18-(10.4)^2}=\mathbf{1.87}$$

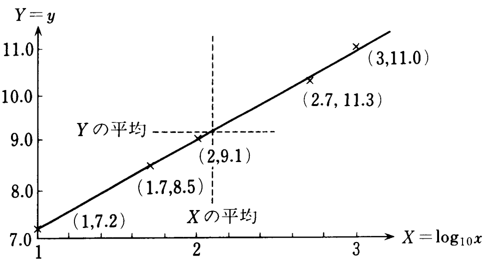
$$a=\overline{Y}-b\overline{X}$$

$$=9.22-1.87\times2.08$$

$$=\mathbf{5.33}$$

よって，Y の X への回帰直線は

$$Y=5.33+1.87X$$



例題 4（相関係数，回帰直線の計算）

10人の生徒について，身長とひじの長さを測定して次のデータを得た(単位：cm)．

身長(x)	152	145	164	153	132	148	149	138	142	150
ひじの長さ(y)	22	21	25	23	20	21	22	21	22	24

身長とひじの長さの相関係数を求めよ．また，Y の X への回帰直線を求めよ．身長が 150 cm の生徒のひじの長さはどのくらいか．

解　$x-150=u$, $y-20=v$ とおき，

(x と y の相関係数)＝(u と v の相関係数)

(y の x への回帰係数)＝(v の u への回帰係数)

を利用して，計算する．

x	y	u	v	u^2	v^2	uv
152	22	2	2	4	4	4
145	21	−5	1	25	1	−5
164	25	14	5	196	25	70
153	23	3	3	9	9	9
132	20	−18	0	324	0	0
148	21	−2	1	4	1	−2
149	22	−1	2	1	4	−2
138	21	−12	1	144	1	−12
142	22	−8	2	64	4	−16
150	24	0	4	0	16	0
計		−27	21	771	65	46

$$n=10,\quad \bar{u}=-2.7,\quad \bar{v}=2.1$$

$$\therefore\quad \bar{x}=\bar{u}+150=147.3$$

$$\bar{y}=\bar{v}+20=22.1$$

相関係数は

$$r_{xy}=r_{uv}=\frac{n\sum u_iv_i-(\sum u_i)(\sum v_i)}{\sqrt{\{n\sum u_i{}^2-(\sum u_i)^2\}}\sqrt{\{n\sum v_i{}^2-(\sum v_i)^2\}}}$$

$$=\frac{10\times 46-(-27)\times 21}{\sqrt{\{10\times 771-(-27)^2\}}\sqrt{\{10\times 65-(21)^2\}}}=\mathbf{0.85}$$

回帰係数は

$$b=\frac{n\sum u_iv_i-(\sum u_i)(\sum v_i)}{n\sum u_i^2-(\sum u_i)^2}=\frac{10\times 46-(-27)\times 21}{10\times 771-(-27)^2}=\mathbf{0.147}$$

よって，Y の X への回帰直線は

$$y-22.1=0.147(x-147.3)$$

$$\therefore\quad \boldsymbol{y=0.45+0.147x}$$

$x=150$ に対する y の値は

$$y=0.45+0.147\times 150=\mathbf{22.5}$$

例題 5　（順位相関係数）

（a）　出品された 8 個の作品を 2 人の審査員が独立に評価して，次の順位を与えた．スピアマンの順位相関係数を求めよ．

作品の番号	1	2	3	4	5	6	7	8
審査員 A	7	4	3	1	6	8	5	2
審査員 B	5	8	1	6	2	7	3	4

（b）　6 人の学生の数学と英語の成績が 5 段階評価で次のように与えられた．数学の成績と英語の成績のスピアマン順位相関係数を求めよ．

学生の番号	1	2	3	4	5	6
数学（X）	A	B	B	E	D	C
英語（Y）	C	B	A	E	E	B

解　（a）

A	B	$d=A-B$	d^2
7	5	2	4
4	8	-4	16
3	1	2	4
1	6	-5	25
6	2	4	16
8	7	1	1
5	3	2	4
2	4	-2	4
計		0	74

順位相関係数の公式から

$$r_S=1-\frac{6\sum d_i{}^2}{n(n^2-1)}=1-\frac{6\times 74}{8(8^2-1)}=\mathbf{0.12}$$

（b）　数学では，学生番号の 2 番と 3 番が同順位なので，平均順位 2.5 を与え，英語では，2 番と 6 番および 4 番と 5 番が同順位なので，前者には平均順位 2.5 を与え，後者には平均順位 5.5 を与える．

X	Y	$d=X-Y$	d^2
1	4	-3	9
2.5	2.5	0	0
2.5	1	1.5	2.25
6	5.5	0.5	0.25
5	5.5	-0.5	0.25
4	2.5	1.5	2.25
計		0	14.0

$$r_S=1-\frac{6\times 14.0}{6(6^2-1)}=\mathbf{0.6}$$

例題 6　（相関係数の検定）

A 地区のどぶネズミの体重と尾の長さの相関係数は，これまでの観測から，$\rho=0.75$ であることが知られている．いま B 地区の 12 匹のどぶネズミの体重と尾の長さの標本相関係数を求めたら，$r=0.62$ となった．この地区のどぶネズミの体重と尾の長さの相関係数は A 地区のそれと同じであるといえるか．

解　帰無仮説を $H_0:\rho=0.75$，対立仮説を $H_1:\rho\neq 0.75$ とし，有意水準 $\alpha=0.05$ の両側検定を行う．検定統計量

$$z=\frac{1}{2}\log_e\frac{1+r}{1-r}\qquad (r\text{ は標本相関係数})$$

は，H_0 の下で $z\sim N\left(\frac{1}{2}\log_e\frac{1+\rho}{1-\rho},\ \frac{1}{n-3}\right)$ であるから，$\alpha=0.05$ の棄却域は

$$\frac{\left|z-\frac{1}{2}\log_e\frac{1+\rho}{1-\rho}\right|}{\frac{1}{\sqrt{n-3}}}>1.96$$

$n=12, \rho=0.75$ を代入すれば，この棄却域は

$$z<0.32 \quad \text{または} \quad z>1.63$$

$r=0.62$ に対する $z=\dfrac{1}{2}\log_e\dfrac{1+0.62}{1-0.62}=0.725$ は棄却域に入らないから，H_0 は棄却されない．したがって B 地区の相関係数は，A 地区と比べて**差があるとはいえない**．

例題 7　**（予測値の信頼限界）**

スプリングの変位を測定して次のデータを得た．

荷重（x）	0	10	20	30	40
延び（y）	18.2	22.3	27.0	31.3	34.2

（a）　Y の X への回帰直線を求めよ．この直線を散布図上にかき入れよ．

（b）　$x=35$ のとき，y の値に対する 95%信頼区間を求めよ．

解　データから

	x	y	$x-\bar{x}$	$y-\bar{y}$	$(x-\bar{x})^2$	$(y-\bar{y})^2$	$(x-\bar{x})(y-\bar{y})$
	0	18.2	-20	-8.4	400	70.56	168
	10	22.3	-10	-4.3	100	18.49	43
	20	27.0	0	0.4	0	0.16	0
	30	31.3	10	4.7	100	22.09	47
	40	34.2	20	7.6	400	57.76	152
計	100	133	0	0	1000	169.06	410

（a）　$n=5,\ \bar{x}=20,\ \bar{y}=26.6,\ \sum(x_i-\bar{x})^2=1000,\ \sum(y_i-\bar{y})^2=169.06,$
$\sum(x_i-\bar{x})(y_i-\bar{y})=410$

であるから回帰係数は

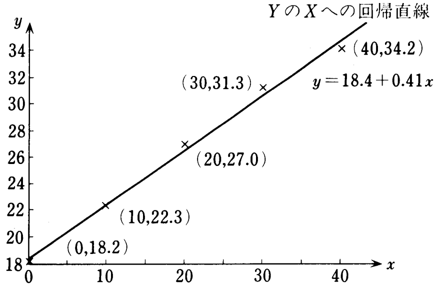
$$b=\frac{\sum(x_i-\bar{x})(y_i-\bar{y})}{\sum(x_i-\bar{x})^2}=\frac{410}{1000}=0.41$$

よって Y の X への回帰直線は

$$y=\bar{y}+b(x-\bar{x})=26.6+0.41(x-20)$$

すなわち，

$$\boldsymbol{y=18.4+0.41x}$$



（b）　σ^2 の推定値 $s_e{}^2$ は

$$s_e{}^2=\frac{1}{n-2}\{\sum(y_i-\bar{y})^2-b\sum(x_i-\bar{x})(y_i-\bar{y})\}$$

$$=\frac{1}{3}\{169.06-0.41\times 410\}=0.32$$

ゆえに，$s_e=0.566$.

$x=35$ に対する y の予測値の95%信頼限界は，$n=5$，$x_0=35$，$t_{0.025}(3)=3.182$ より

$$a+bx_0\pm t_{\frac{\alpha}{2}}(n-2)s_e\sqrt{\frac{1}{n}+\frac{(x_0-\bar{x})^2}{\sum(x_i-\bar{x})^2}}$$

$$=18.4+0.41\times 35\pm 3.182\times 0.566\sqrt{\frac{1}{5}+\frac{(35-20)^2}{1000}}$$

$$=32.75\pm 1.17$$

$$=31.58,\ 33.92$$

よって求める信頼区間は **(31.58, 33.92)**.

例題 8　（最小 2 乗法）

ある微生物の成長モデルは時間 x と大きさ y の関係が

$$y=Ae^{Bx}$$

であるという．いま x と y について 5 個の観測値

x	1	2	3	4	5
y	5.5	14.9	39.9	110.2	300.1

を得たとき，A，B の値を最小 2 乗法で求めよ．

解 モデルの式で両辺の対数をとれば，$\log y=\log A+Bx$ となる．$Y=\log y$，$X=x$，$a=\log A$，$b=B$ とすれば，1次式

$$Y=a+bX$$

を得る．最小2乗法によって a と b を推定する．データから

$X=x$	1	2	3	4	5
$Y=\log y$	1.70	2.70	3.69	4.70	5.70

であるから，

$$n=5,\quad \sum X_i=15,\quad \sum Y_i=18.49,\quad \sum X_iY_i=65.47,\quad \sum X_i^2=55$$

したがって

$$b=\frac{5\times 65.47-15\times 18.49}{5\times 55-15^2}=1.0$$

$$a=\overline{Y}-b\overline{X}=3.7-1.0\times 3=0.7$$

よって求める1次式は，$Y=0.7+X$．

$$A=e^{0.7}=\mathbf{2.0},\qquad B=b=\mathbf{1.0}$$

ゆえに，この成長モデルの式は

$$y=2e^x$$

例題 9 （2つの回帰直線）

8人の男子学生の身長（X）と右足（Y）の大きさを測って，次の値を得た．

身長（x cm）	174	156	165	178	164	162	166	171
右足の大きさ（y cm）	24.5	22.5	25.0	26.5	23.0	24.0	25.0	26.0

このデータより，Y の X への回帰直線と X の Y への回帰直線を求めよ．

適切な回帰直線によって次を求めよ．

（a） 身長が 168 cm の男子学生の右足の大きさ．

（b） 右足の大きさが 25.5 cm の男子学生の身長．

解 計算の簡略化のため，$x-165=u$，$y-25=v$ とおく．両方の回帰直線を求めるには $\sum u_i$，$\sum v_i$，$\sum u_i^2$，$\sum v_i^2$，$\sum u_iv_i$ の値が必要である．

x	y	u	v	u^2	v^2	uv
174	24.5	9	−0.5	81	0.25	−4.5
156	22.5	−9	−2.5	81	6.25	22.5
165	25.0	0	0	0	0	0
178	26.5	13	1.5	169	2.25	19.5
164	23.0	−1	−2.0	1	4.0	2
162	24.0	−3	−1.0	9	1.0	3
166	25.0	1	0	1	0	0
171	26.0	6	1.0	36	1.0	6
計		16	−3.5	378	14.75	48.5

よって，$n=8$，$\sum u_i=16$，$\sum v_i=-3.5$，$\sum u_i{}^2=378$，$\sum v_i{}^2=14.75$，
$\sum u_iv_i=48.5$

(x, y)から(u, v)へ変換しても回帰係数は変らないから，Y の X への回帰係数を b，X の Y への回帰係数を b' とすれば

$$b=\frac{n\sum u_iv_i-(\sum u_i)(\sum v_i)}{n\sum u_i{}^2-(\sum u_i)^2}=\frac{8\times48.5-16\times(-3.5)}{8\times378-16^2}=0.16$$

$$b'=\frac{n\sum u_iv_i-(\sum u_i)(\sum v_i)}{n\sum v_i{}^2-(\sum v_i)^2}=\frac{8\times48.5-16\times(-3.5)}{8\times14.75-(-3.5)^2}=4.20$$

$$\bar{x}=165+\bar{u}=165+\frac{16}{8}=167$$

$$\bar{y}=25+\bar{v}=25+\frac{3.5}{8}=24.5625$$

よって，Y の X への回帰直線は

$$y-24.56=0.16(x-167)$$

$$\boldsymbol{y=0.16x-2.16} \qquad (1)$$

同様にして，X の Y への回帰直線は

$$\boldsymbol{x=4.20\,y+63.8} \qquad (2)$$

（a）　(1)より，$x=168$ に対する y の値は

$$y=0.16\times168-2.16=\mathbf{24.72}$$

（b）　(2)より，$y=25.5$ に対する x の値は

$$x=4.2\times25.5+63.8=\mathbf{170.9}$$

10章の問題

10.1 次の表は10人の生徒の化学と物理の得点を示す．化学と物理の両得点の間の相関係数 r と順位相関係数 r_S を求めよ．

	生徒									
	A	B	C	D	E	F	G	H	I	J
化学（x）	83	91	79	77	71	68	43	46	30	39
物理（y）	77	89	95	71	59	55	59	43	35	48

10.2 2つの離散変量 X と Y は，それぞれ2通りの可能な値 $\{p, q\}$ と $\{r, s\}$ をとるとする．いま，(X, Y) についての n 個のデータが，次のように2×2分割表で与えられるとき，X と Y の相関係数を求めよ．

		x の値		計
		p	q	
y の値	r	a	b	$a+b$
	s	c	d	$c+d$
	計	$a+c$	$b+d$	$n=a+b+c+d$

10.3 A，B 2人のOLに5種類のオーディコロンの試作品を与えて，100点満点で点数をつけてもらった．

	①	②	③	④	⑤
A	50	80	75	70	60
B	60	100	100	80	50

この2人の好みは，かなり近いといえるだろうか．

10.4 2変量のデータ (x_i, y_i), $i=1, 2, \cdots, n$ が，順位データであるときの相関係数はスピアマンの順位相関係数

$$r_S = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2-1)}$$

になることを示せ．ただし $d_i = x_i - y_i$ とする．

10.5 美人コンテストで3人の審査員が8人の出場者に次の順位をつけた.

	出場者							
	A	B	C	D	E	F	G	H
審査員1	2	5	6	7	8	1	4	3
審査員2	5	1	4	3	6	2	7	8
審査員3	8	5	6	4	7	3	2	1

各審査員の間の順位相関係数を求めよ.

あなたが主任審査員であるとすれば, このコンテストの第1位, 第2位, 第3位を誰にするか.

10.6 国語 (X) と英語 (Y) の試験を10人の生徒に行ったところ, 次のデータを得た.

生徒	1	2	3	4	5	6	7	8	9	10
国語	39	62	68	61	57	74	41	47	43	65
英語	43	71	67	67	70	89	46	69	58	76

(X, Y) は正規分布に従うとして, 帰無仮説 $H_0: \rho=0$ を対立仮説 $H_1: \rho\neq 0$ に対して1%有意水準で検定せよ.

10.7 理科の実験で, 針金からバネばかりを作った. これを用いておもりをいろいろ変えたときのバネの長さを測定し, 次のデータを得た.

おもり (x g)	10	12	14	16	18	20
バネの長さ (y cm)	20.2	21.3	21.9	23.0	24.2	25.1

(a) Y の X への回帰直線を求めよ.

(b) 15 g のおもりをのせたら, バネの長さは何 cm になるだろうか.

10.8 ある化学実験での反応は, 温度の1次式であるという. いま3水準の温度で2回ずつ6回測定を行ったときの反応値を示す.

温度 X	20	25	30
反応 Y	140.2	159.7	182.0
	141.8	160.9	179.8

これらのデータを用いて

（a） Y の X への回帰直線を求めよ．

（b） 回帰係数 β の 95%信頼区間を求めよ．

10.9 次のデータの散布図をかけ．

x	1	2	3	4	5	6
y	0.51	0.79	0.94	0.95	0.97	0.99

さらに，(x, y) を

$$\begin{cases} X = e^{-x} \\ Y = y^{-1} \end{cases}$$

によって変換し，その結果を (X, Y) 平面に打点せよ．最小 2 乗法で Y の X への回帰直線を求めよ．また，y を x の関数として表せ．

10.10 次のデータの散布図をかけ．

x	3	8	5	3	8
y	7	2	6	6	3

点 $(\bar{x}, \bar{y})$ を散布図の上に打点せよ．目測によってこれらの点を通る直線を引け．Y の X への回帰直線と X の Y への回帰直線を求めて，散布図の上にかき入れよ．適当な回帰直線により，$y=5$ に対する x の値を推定せよ．

問題解答

1　章

1.1

	(a)	(b)	(c)	(d)
階級	20-29	150-, 180-	1.6-3.5	(−8)-(−5)
真の級限界	19.5-29.5	149.5-179.5	1.55-3.55	(−8.5)-(−4.5)
階級値	24.5	164.5	2.55	−6.5
階級幅	10	30	2	4

1.2　(a)　3, 5, 4, 1, 7 の平均と標準偏差は

$$平均=\frac{3+5+4+1+7}{5}=4$$

$$標準偏差=\sqrt{\frac{(3-4)^2+(5-4)^2+(4-4)^2+(1-4)^2+(7-4)^2}{5}}=2$$

(i)　これは元のデータに定数 10 を定したものである．定数を足すと，平均はその分だけ増えるが，標準偏差は変わらないから

$$平均=\mathbf{14},\quad 標準偏差=\mathbf{2}$$

(ii)　これは元のデータから定数 5 を引いたものである．定数を引くと，平均はその分だけ減るが，標準偏差は変わらないから

$$平均=\mathbf{-1},\quad 標準偏差=\mathbf{2}$$

(iii)　これは元のデータに定数 a を足したものであるから，

$$平均=\boldsymbol{a+4},\quad 標準偏差=\mathbf{2}$$

(iv)　これは元のデータに定数 $\frac{1}{10}$ を掛けたものである．定数を掛けると，平均平均も標準偏差もその定数倍されるから

$$平均=\mathbf{0.4},\quad 標準偏差=\mathbf{0.2}$$

(v)　これも元のデータに定数 a を掛けて，定数 b を加えたものであるから

$$平均=\boldsymbol{4a+b},\quad 標準偏差=\boldsymbol{2a}$$

(b)　$平均=\dfrac{\boldsymbol{5a+6b+4c+3d+2e}}{\mathbf{20}}$

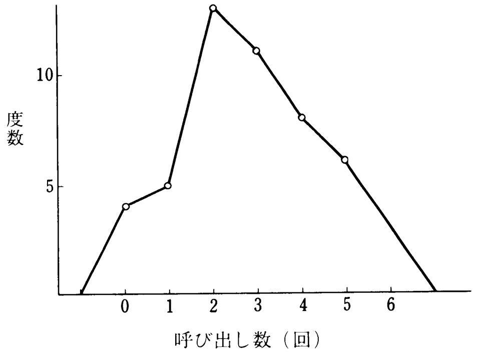
小さい方から 10 番目と 11 番目はいずれも b であるから，メジアンは $\boldsymbol{b}$.
b が 6 個で最も多く出ているから，モードも $\boldsymbol{b}$.

1.3

度数分布

x	f	F	xf
0	4	4	0
1	5	9	5
2	13	22	26
3	11	33	33
4	8	41	32
5	6	47	30
6	3	50	18
計	50		144

度数多角形



平均値は $\dfrac{144}{50}$≒**2.9**. 累積度数 F の値より 25 番目と 26 番目はともに 3 だから，メジアンは **3**. 最大度数は 13 だから，モードは **2**.

1.4

階級	真の級限界	階級値(x)	検数マーク	度数(f)	累積度数(F)
55- 64	54.5- 64.5	59.5	一	1	1
65- 74	64.5- 74.5	69.5	下	3	4
75- 84	74.5- 84.5	79.5	正正下	13	17
85- 94	84.5- 94.5	89.5	正正正under	19	36
95-104	94.5-104.5	99.5	正正正正正𠄌	29	65
105-114	104.5-114.5	109.5	正正正正丅	22	87
115-124	114.5-124.5	119.5	正正	10	97
125-134	124.5-134.5	129.5	丅	2	99
135-144	134.5-144.5	139.5	一	1	100
計				100	

よって，

（a）

x	59.5	69.5	79.5	89.5	99.5	109.5	119.5	129.5	139.5	計
f	1	3	13	19	29	22	10	2	1	100

(b)

x	f	$u=\dfrac{x-99.5}{10}$	uf	u^2f
59.5	1	−4	−4	16
69.5	3	−3	−9	27
79.5	13	−2	−26	52
89.5	19	−1	−19	19
99.5	29	0	0	0
109.5	22	1	22	22
119.5	10	2	20	40
129.5	2	3	6	18
139.5	1	4	4	16
計	100		−6	210

累積度数分布表

真の級限界の上限	累積度数
64.5	1
74.5	4
89.5	17
94.5	36
m	50
104.5	65
114.5	87
124.5	97
134.5	99
144.5	100

$$\bar{u}=\frac{-6}{100}=-0.06$$

$$\bar{x}=99.5+10\times(-0.06)\fallingdotseq\mathbf{98.9}$$

$$s=10\sqrt{\frac{210}{100}-(-0.06)^2}\fallingdotseq\mathbf{14.5}$$

(c) メジアンを m とすると，右上の累積度数分布表より

$$\frac{m-94.5}{104.5-94.5}=\frac{50-36}{65-36} \quad\Rightarrow\quad m=94.5+10\times\frac{14}{29}\fallingdotseq\mathbf{99.3}$$

1.5 平均値 $=\dfrac{x+5+y+13}{4}=7 \quad\Longrightarrow\quad x+y=10$

メジアン $=\dfrac{5+y}{2}=6 \quad\Longrightarrow\quad y=7$

よって，$x=3, y=7$. したがって分散は

$$s^2=\frac{3^2+5^2+7^2+13^2}{4}-7^2=\mathbf{14}$$

1.6 平均値は

$$\bar{x}=\frac{6\times220+247+250+239+264+5\times280}{15}=\mathbf{248}\ \text{(g)}$$

メジアンは小さい方から 8 番目だから，**247** (g).
標準偏差は

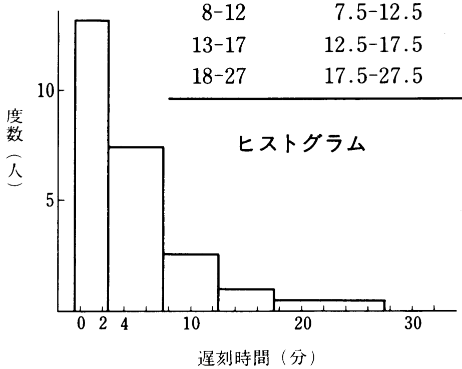
$$s=\sqrt{\frac{20000}{3}-\left(\frac{240}{3}\right)^2}=\mathbf{16.3}\ \text{(g)}$$

1.7 $\bar{x}=\dfrac{2\times2+3\times4+\cdots+12\times2}{100}=\dfrac{706}{100}=\mathbf{7.06}$

$$s=\sqrt{\frac{2^2\times2+3^2\times4+\cdots+12^2\times2}{100}-7.06^2}\fallingdotseq\mathbf{2.3}$$

1.8 (a) 階級の幅が等しくないことに注意.

階級	階級の真の限界	階級の幅	度数	度数密度
0- 2	−0.5- 2.5	3	40	13.3
3- 7	2.5- 7.5	5	37	7.4
8-12	7.5-12.5	5	13	2.6
13-17	12.5-17.5	5	5	1.0
18-27	17.5-27.5	10	5	0.5



(b)

x	f	xf	x^2f
1.0	40	40	40
5.0	37	185	925
10.0	13	130	1300
15.0	5	75	1125
22.5	5	112.5	2531.25
計	100	542.5	5921.25

$$\bar{x}=\frac{542.5}{100}\fallingdotseq\mathbf{5.4}\ (分)$$

$$s=\sqrt{\frac{5921.25}{100}-5.4^2}\fallingdotseq\mathbf{5.5}\ (分)$$

1.9

階級	真の限界値	階級値 x	度数 f	xf	x^2f
1- 20	0.5- 20.5	10.5	4	42.0	441.0
21- 30	20.5- 30.5	25.5	15	382.5	9753.75
31- 40	30.5- 40.5	35.5	43	1526.5	54190.75
41- 50	40.5- 50.5	45.5	55	2502.5	113863.75
51- 60	50.5- 60.5	55.5	39	2164.5	120129.75
61- 70	60.5- 70.5	65.5	31	2030.5	132997.75
71- 80	70.5- 80.5	75.5	10	755.0	57002.5
81-100	80.5-100.5	90.5	3	271.5	24570.75
		計	200	9675.0	512950.00

よって

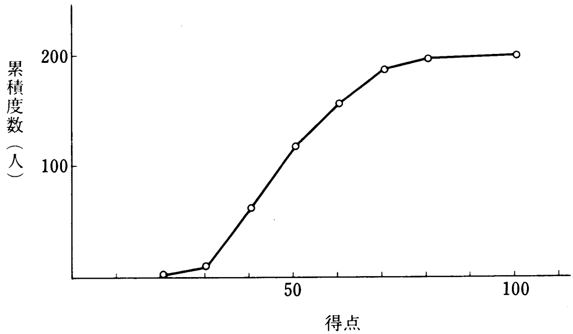
$$\bar{x}=\frac{9675}{200}\fallingdotseq 48.4$$

$$s=\sqrt{\frac{512950.00}{200}-48.4^2}\fallingdotseq 15.0$$

（b） **累積度数分布表**

x	F
20.5	4
30.5	19
40.5	62
50.5	117
60.5	156
70.5	187
80.5	197
100.5	200

累積度数多角形



1.10 10 個の観測値の平均を $\bar{x}$ とすると

$$\bar{x}=\frac{4\times 3.13+3.19+2.86+\cdots+3.21}{10}=\mathbf{3.1}$$

次に，4 個の観測値を x_1, x_2, x_3, x_4, 平均と標準偏差を $\bar{x}_1, s_1$ とすると，公式 $s_1{}^2=\dfrac{\sum_{i=1}^{4}x_i{}^2}{n}-\bar{x}_1{}^2$ より

$$0.15^2=\frac{\sum_{i=1}^{4}x_i{}^2}{4}-3.13^2 \implies \sum_{i=1}^{4}x_i{}^2=39.2776$$

よって，10 個の観測値の標準偏差を s とすれば

$$s=\sqrt{\frac{39.2776+3.19^2+2.86^2+\cdots+3.21^2}{10}-3.1^2}\fallingdotseq\mathbf{0.14}$$

1.11 （a） $n_1=4$, $n_2=6$

$\bar{x}_1=5$, $\bar{x}_2=7$

$s_1{}^2=2$, $s_2{}^2=3$

公式 $\sum x_i{}^2=n(\bar{x}^2+s^2)$ より

最初の 4 個の数の 2 乗和は $\sum_{i=1}^{4}x_i{}^2=4(5^2+2)=108$

次の 6 個の数の 2 乗和は $\sum_{i=1}^{6}x_i{}^2=6(7^2+3)=312$

よって，合併データの平均と分散は

$$\bar{x}=\frac{4\times 5+6\times 7}{10}=\mathbf{6.2}$$

$$s^2=\frac{108+312}{10}-6.2^2=\mathbf{3.56}$$

2 章

2.1 (a) (i) $P(A)=1-P(\overline{A})=1-\dfrac{5}{8}=\mathbf{\dfrac{3}{8}}$

(ii) $\dfrac{7}{8}=P(A\cup B)=P(A)+P(B)-P(A\cap B)$

$$=\frac{3}{8}+P(B)-\frac{1}{4}\Rightarrow P(B)=\frac{3}{4}\Rightarrow P(\overline{B})=\mathbf{\frac{1}{4}}$$

(iii) $P(A\cap\overline{B})=P(A)-P(A\cap B)=\dfrac{3}{8}-\dfrac{1}{4}=\mathbf{\dfrac{1}{8}}$

(b) (i) $P(A\,|\,B)=\dfrac{P(A\cap B)}{P(B)}=\dfrac{1/5}{1/3}=\mathbf{\dfrac{3}{5}}$

(ii) $P(B\,|\,A)=\dfrac{P(A\cap B)}{P(A)}=\dfrac{1/5}{1/2}=\mathbf{\dfrac{2}{5}}$

(iii) $P(A\cup B)=P(A)+P(B)-P(A\cap B)=\dfrac{1}{2}+\dfrac{1}{3}-\dfrac{1}{5}=\mathbf{\dfrac{19}{30}}$

(iv) $P(\overline{A}|B)=1-P(A|B)=1-\dfrac{3}{5}=\mathbf{\dfrac{2}{5}}$

2.2 (a) $P(X)=\mathbf{\dfrac{1}{3}}$

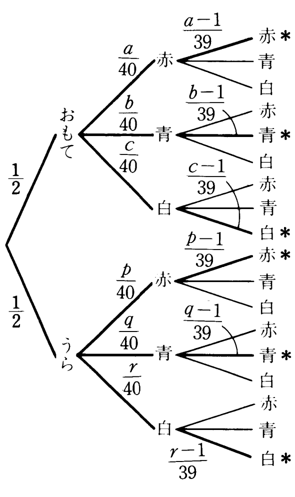
(b) $P(Y)=\dfrac{1}{3}\times\dfrac{2}{10}+\dfrac{2}{3}\times\dfrac{3}{10}=\mathbf{\dfrac{4}{15}}$

(c) $P(X\cap Y)=\dfrac{1}{3}\times\dfrac{2}{10}=\mathbf{\dfrac{1}{15}}$

(d) $P(Y\,|\,X)=\dfrac{P(X\cap Y)}{P(X)}=\dfrac{1/15}{1/3}=\mathbf{\dfrac{1}{5}}$

(e) $P(\overline{X}\cap Y)=P(Y)-P(X\cap Y)=\dfrac{4}{15}-\dfrac{1}{15}=\mathbf{\dfrac{1}{5}}$

2.3 樹形図をかくと



$$a+b+c=p+q+r=40$$

樹形図より，同じ色の球がとり出されるのは，＊印で示した 6 つの場合であるから，求める確率 P は

$$P=\frac{1}{2}\left(\frac{a}{40}\cdot\frac{a-1}{39}+\frac{b}{40}\cdot\frac{b-1}{39}+\frac{c}{40}\cdot\frac{c-1}{39}\right)$$

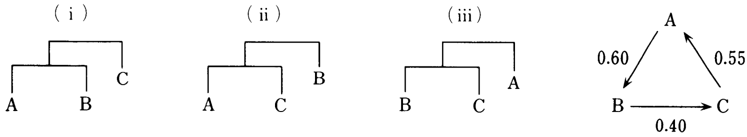
$$+\frac{1}{2}\left(\frac{p}{40}\cdot\frac{p-1}{39}+\frac{q}{40}\cdot\frac{q-1}{39}+\frac{r}{40}\cdot\frac{r-1}{39}\right)$$

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

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

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

2.4 A の優勝は次の 3 つの場合で起こる.



これら 3 つの場合は互いに排反で，同じ確率で起こる.

A が (i) で優勝する確率は，A がまず B に勝ち，その後 C に勝てばよいから，

$$\frac{1}{3}\times 0.60\times 0.45=0.09$$

同様にして (ii) で，優勝する確率は $\frac{1}{3}\times 0.45\times 0.60=0.09$.

A が (iii) で優勝する確率は $\frac{1}{3}(0.40\times 0.60+0.60\times 0.45)=0.17$.

よって，Aが (i) または (ii) または (iii) で優勝する確率は

$$0.09+0.09+0.17=\mathbf{0.35}.$$

2.5 P(A が勝つ)$=P$(A が 1 番目に白球をとって勝つ)

$+P$(A が 4 番目に初めて白球をとって勝つ)

$$=\frac{5}{10}+\frac{5}{10}\times\frac{4}{9}\times\frac{3}{8}\times\frac{5}{7}=\mathbf{\frac{47}{84}}$$

P(B が勝つ)$=P$(B が 2 番目に初めて白球をとって勝つ)

$+P$(B が 5 番目に初めて白球をとって勝つ)

$$=\frac{5}{10}\times\frac{5}{9}+\frac{5}{10}\times\frac{4}{9}\times\frac{3}{8}\times\frac{2}{7}\times\frac{5}{6}=\mathbf{\frac{25}{84}}$$

P(C が勝つ)$=P$(C が 3 番目に初めて白球をとって勝つ)

$+P$(C が 6 番目に初めて白球をとって勝つ)

$$=\frac{5}{10}\times\frac{4}{9}\times\frac{5}{8}+\frac{5}{10}\times\frac{4}{9}\times\frac{3}{8}\times\frac{2}{7}\times\frac{1}{6}\times\frac{5}{5}=\mathbf{\frac{1}{7}}$$

2.6 (a) (i) $\dfrac{4\times{}_{13}C_3}{{}_{52}C_3}=\mathbf{\dfrac{22}{425}}$, (ii) $\dfrac{13\times{}_4C_3}{{}_{52}C_3}=\mathbf{\dfrac{1}{425}}$,

(iii) $\dfrac{22}{425}+\dfrac{1}{425}=\mathbf{\dfrac{23}{425}}$

(b) (i)

目の和が 8 になる場合.		目の和が 9 になる場合.	
(1,1,6)	順列は 3 通り	(1,2,6)	順列は 6 通り
(1,2,5)	順列は 6 通り	(1,3,5)	順列は 6 通り
(1,3,4)	順列は 6 通り	(2,2,5)	順列は 3 通り
(2,2,4)	順列は 3 通り	(1,4,4)	順列は 3 通り
(2,3,3)	順列は 3 通り	(2,3,4)	順列は 6 通り
		(3,3,3)	順列は 1 通り
計 21 通り		計 25 通り	

よって，P(目の和が 8 または 9)$=\dfrac{21+25}{216}=\mathbf{\dfrac{23}{108}}$.

（ii） 目の和が完全平方になるのは，4，9，16のときで，4になるのは(1,1,2)の3通り，9になるのは(i)より25通り，16になるのは，(5,5,6)の3通りと(4,6,6)の3通りの計6通りだから，全部で，$3+25+6=34$通り．よって，P(目の和が完全平方)$=\dfrac{34}{216}=\mathbf{\dfrac{17}{108}}$．

2.7 最初にとり出した球は元にもどすから，最初の結果と2回目の結果は独立である．いま，赤球をR，白球をWとすると，

（a） $P\{(\mathrm{R,R})\cap(\mathrm{W,W})\}$

$$=P(\mathrm{R,R})\cdot P(\mathrm{W,W})=\frac{{}_3\mathrm{C}_2}{{}_7\mathrm{C}_2}\cdot\frac{{}_4\mathrm{C}_2}{{}_7\mathrm{C}_2}=\frac{1}{7}\times\frac{2}{7}=\mathbf{\frac{2}{49}}$$

（b） これは，最初が(R, R)で，次が(W, W)か，
最初が(W, W)で，次が(R, R)か，
最初が(R, W)で，次も(R, W)

のいずれかの場合で起こる．初めの2つの場合の確率はどちらも $\dfrac{2}{49}$ ((a)より)で，第3の場合の確率は

$$P\{(\mathrm{R,W})\cap(\mathrm{R,W})\}=P(\mathrm{R,W})\cdot P(\mathrm{R,W})=\left(\frac{{}_3\mathrm{C}_1\cdot{}_4\mathrm{C}_1}{{}_7\mathrm{C}_2}\right)=\frac{16}{49}$$

となるから，求める確率は

$$\frac{2}{49}+\frac{2}{49}+\frac{16}{49}=\mathbf{\frac{20}{49}}$$

（c） 4個とも赤球が得られる確率は

$$P(\mathrm{R,R})\cdot P(\mathrm{R,R})=\left(\frac{{}_3\mathrm{C}_1}{{}_7\mathrm{C}_2}\right)^2=\frac{1}{49}$$

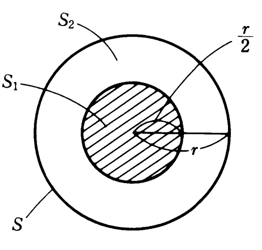
同様に，4個とも白球が得られる確率は

$$P(\mathrm{W,W})\cdot P(\mathrm{W,W})=\left(\frac{{}_4\mathrm{C}_1}{{}_7\mathrm{C}_2}\right)^2=\frac{4}{49}$$

よって，求める確率は

$$\frac{1}{49}+\frac{4}{49}=\mathbf{\frac{5}{49}}$$

2.8 円の半径を r とする．ランダムに選んだ点が円周より中心に近い方に入るのは，その点が半径 $\dfrac{1}{2}r$ の円内にあるときだから，求める確率は2つの円の面積の比で与えられる．



$$\frac{\pi\left(\frac{1}{2}r\right)^2}{\pi r^2}=\frac{1}{4}$$

（a） ドーナツ部分を S_2 とすると，最初の3点が S_2 に入り，4番目の点が S_1 に入る確率を求めればよいから

$$\left(1-\frac{1}{4}\right)^3\times\frac{1}{4}=\mathbf{\frac{27}{256}}$$

（b） 1, 2, 3, ･･･, n 回目のいずれかで初めて S_1 内に入る確率が 0.90 よりも大きくなる n を求めればよいから

$$\sum_{i=1}^{n}\left(\frac{3}{4}\right)^{n-1}\cdot\frac{1}{4}\geqq 0.90$$

これから

$$\frac{\frac{1}{4}\left\{1-\left(\frac{3}{4}\right)^{n}\right\}}{1-\frac{3}{4}}\geqq 0.90$$

よって

$$\left(\frac{3}{4}\right)^{n}\leqq 0.10$$

$$\therefore\quad n\geqq\frac{\log 0.10}{\log 0.75}\fallingdotseq 8.01$$

よって，**9 個**以上を必要とする．

2.9 （a） $0.35\times 0.30=\mathbf{0.105}$

（b） $0.80\times 0.30\times 0.60+0.20\times 0.35\times 0.70=\mathbf{0.193}$

（c） $\dfrac{0.80\times 0.30\times 0.60\times 0.20\times 0.35\times 0.70}{0.80\times 0.60+0.20\times 0.70}\fallingdotseq\mathbf{0.311}$

2.10 $\dfrac{0.01\times 0.90}{0.01\times 0.90+0.99\times 0.005}\fallingdotseq\mathbf{0.645}$

3 章

3.1 （a） $E(X)=0\times 0.4+2\times 0.1+5\times 0.2+9\times 0.3=\mathbf{3.9}$

$V(X)=0^2\times 0.4+2^2\times 0.1+5^2\times 0.2+9^2\times 0.3-3.9^2=\mathbf{14.49}$

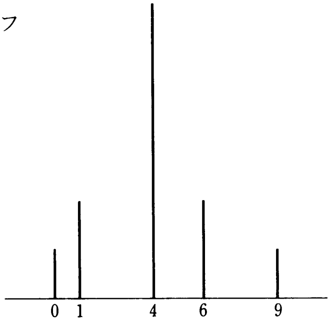
（b） $E(X)=\mathbf{0}$,

$V(X)=(-2)^2\times 0.1+(-1)^2\times 0.2+0^2\times 0.4+1^2\times 0.2+2^2\times 0.1-0^2$
$=\mathbf{1.2}$

（c） $E(X)=(-2)\times\dfrac{1}{8}+2\times\dfrac{1}{4}+4\times\dfrac{3}{8}+5\times\dfrac{1}{4}=\mathbf{3}$

$V(X)=(-2)^2\times\dfrac{1}{8}+2^2\times\dfrac{1}{4}+4^2\times\dfrac{3}{8}+5^2\times\dfrac{1}{4}-3^2=\mathbf{\dfrac{19}{4}}$

3.2 確率分布のグラフ



$$E(X)=0\times\frac{1}{12}+1\times\frac{1}{6}+4\times\frac{1}{2}+6\times\frac{1}{6}+9\times\frac{1}{12}=\mathbf{\frac{47}{12}}$$

$$V(X)=0^2\times\frac{1}{12}+1^2\times\frac{1}{6}+4^2\times\frac{1}{2}+6^2\times\frac{1}{6}+9^2\times\frac{1}{12}-\left(\frac{47}{12}\right)^2=\mathbf{\frac{803}{144}}$$

$$E(Y)=E(5X+2)=5E(X)+2=5\times\frac{47}{12}+2=\mathbf{\frac{259}{12}}$$

$$V(Y)=V(5X+2)=25V(X)=25\times\frac{803}{144}=\mathbf{\frac{20075}{144}}$$

3.3 (a) $E(X)=0\times0.1+1\times0.1+2\times0.3+3\times0.3+4\times0.2=\mathbf{2.4}$

$V(X)=0^2\times0.1+1^2\times0.1+2^2\times0.3+3^2\times0.3+4^2\times0.2-2.4^2$

$=7.2-5.76=\mathbf{1.44}$

(b) $E(X^2)=\mathbf{7.2}$

$V(X^2)=E(X^4)-\{E(X^2)\}^2$

$=0^4\times0.1+1^4\times0.1+2^4\times0.3+3^4\times0.3+4^4\times0.2-7.2^2$

$=\mathbf{28.56}$

3.4 $p+p^2+2p^2+p=1 \Rightarrow 3p^2+2p-1=0 \Rightarrow (p+1)(3p-1)=0 \Rightarrow p=\frac{1}{3}$

x	0	1	2	3
$P(X=x)$	$\frac{1}{3}$	$\frac{1}{9}$	$\frac{2}{9}$	$\frac{1}{3}$

$$E(X)=0\times\frac{1}{3}+1\times\frac{1}{9}+2\times\frac{2}{9}+3\times\frac{1}{3}=\mathbf{\frac{14}{9}}$$

$$V(X)=0^2\times\frac{1}{3}+1^2\times\frac{1}{9}+2^2\times\frac{2}{9}+3^2\times\frac{1}{3}-\left(\frac{14}{9}\right)^2=\mathbf{\frac{128}{81}}$$

3.5 $E(X)=2,\ V(X)=5$

(a) $E(X-1)=E(X)-1=\mathbf{1}$, $V(X-1)=V(X)=\mathbf{5}$

(b) $E3X)=3E(X)=\mathbf{6}$, $V(3X)=9V(X)=\mathbf{45}$

(c) $E(2X+1)=2E(X)+1=\mathbf{5}$, $V(2X+1)=4V(X)=\mathbf{20}$

(d) $E\left\{\frac{1}{3}(X+3)\right\}=\frac{1}{3}E(X)+1=\mathbf{\frac{5}{3}}$, $V\left\{\frac{1}{3}(X+3)\right\}=\frac{1}{9}V(X)=\mathbf{\frac{5}{9}}$

(e) $E(5-3X)=5-3E(X)=\mathbf{-1}$, $V(5-3X)=9V(X)=\mathbf{45}$

3.6

Xの確率分布：

x	3	4	5	6	7	8	9	10	11
$P(X=x)$	$\frac{1}{15}$	$\frac{1}{15}$	$\frac{2}{15}$	$\frac{2}{15}$	$\frac{3}{15}$	$\frac{2}{15}$	$\frac{2}{15}$	$\frac{1}{15}$	$\frac{1}{15}$

Yの確率分布：

y	2	3	4	5	6	8	10	12	15	18	20	24	30
$\mathrm{P}(Y=y)$	$\frac{1}{15}$	$\frac{1}{15}$	$\frac{1}{15}$	$\frac{1}{15}$	$\frac{2}{15}$	$\frac{1}{15}$	$\frac{1}{15}$	$\frac{2}{15}$	$\frac{1}{15}$	$\frac{1}{15}$	$\frac{1}{15}$	$\frac{1}{15}$	$\frac{1}{15}$

3.7 $P(X=x)=\dfrac{{}_4C_x\ {}_{48}C_{13-x}}{{}_{52}C_{13}}$ $(x=0, 1, 2, 3, 4)$

3.8 (b)． X と Y が独立ならば

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

が成り立つが，X と X は明らかに独立でないから

$$V(X+X)\neq V(X)+V(X)$$

よって，(b)は成り立たない．

3.9 まず，この実験の標本空間 S は

$$S=\left\{\begin{matrix}(1,1), & (1,2), & \cdots, & (1,6)\\ \vdots & \vdots & \ddots & \vdots\\ (6,1), & (6,2), & \cdots, & (6,6)\end{matrix}\right\}$$

(a) これより，Z の確率分布は

z	0	1	2	3	4	5
$P(Z=z)$	$\frac{6}{36}$	$\frac{10}{36}$	$\frac{8}{36}$	$\frac{6}{36}$	$\frac{4}{36}$	$\frac{2}{36}$

平均と分散は

$$E(Z)=0\times\frac{6}{36}+1\times\frac{1}{36}+2\times\frac{8}{36}+3\times\frac{6}{36}+4\times\frac{4}{36}+5\times\frac{2}{36}=\mathbf{\frac{35}{18}}$$

$$V(Z)=0^2\times\frac{6}{36}+1^2\times\frac{1}{36}+2^2\times\frac{8}{36}+3^2\times\frac{6}{36}+4^2\times\frac{4}{36}+5^2\times\frac{2}{36}-\left(\frac{35}{18}\right)^2=\mathbf{\frac{665}{9}}$$

(b) また，W の確率分布は

w	1	2	3	4	5	6
$P(W=w)$	$\frac{11}{36}$	$\frac{9}{36}$	$\frac{7}{36}$	$\frac{5}{36}$	$\frac{3}{36}$	$\frac{1}{36}$

$$E(W)=1\times\frac{11}{36}+2\times\frac{9}{36}+3\times\frac{7}{36}+4\times\frac{5}{36}+5\times\frac{3}{36}+6\times\frac{1}{36}=\mathbf{\frac{91}{36}}$$

$$V(W)=1^2\times\frac{11}{36}+2^2\times\frac{9}{36}+3^2\times\frac{7}{36}+4^2\times\frac{5}{36}+5^2\times\frac{3}{36}+6^2\times\frac{1}{36}-\left(\frac{91}{36}\right)^2=\mathbf{\frac{2555}{36}}$$

3.10 この 2 元表の周辺和をそれぞれ求める．

x \ y	0	1	2	$P(X=x)$
0	$\frac{1}{60}$	$\frac{1}{30}$	$\frac{1}{20}$	$\frac{1}{10}$
1	$\frac{1}{30}$	$\frac{1}{15}$	$\frac{1}{10}$	$\frac{1}{5}$
2	$\frac{1}{20}$	$\frac{1}{10}$	$\frac{3}{20}$	$\frac{3}{10}$
3	$\frac{1}{15}$	$\frac{2}{15}$	$\frac{1}{5}$	$\frac{2}{5}$
$P(Y=y)$	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{2}$	1

よって，

(a)

X の周辺分布

x	0	1	2	3
$P(X=x)$	$\frac{1}{10}$	$\frac{1}{5}$	$\frac{3}{10}$	$\frac{2}{5}$

Y の周辺分布

y	0	1	2
$P(Y=y)$	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{2}$

(b)　2 元表から，

$$P(X=0)P(Y=0)=\frac{1}{10}\times\frac{1}{6}=\frac{1}{60}=P(X=0,\ Y=0)$$

$$P(X=1)P(Y=0)=\frac{1}{5}\times\frac{1}{6}=\frac{1}{30}=P(X=1,\ Y=0)$$

残り 10 通りの組合せについても同様なことがいえるから，**X と Y は独立**である．

(c)　(a)より

$$E(X)=0\times\frac{1}{10}+1\times\frac{1}{5}+2\times\frac{3}{10}+3\times\frac{2}{5}=2$$

$$E(Y)=0\times\frac{1}{6}+1\times\frac{1}{3}+2\times\frac{1}{2}=\frac{4}{3}$$

$$V(X)=0^2\times\frac{1}{10}+1^2\times\frac{1}{5}+2^2\times\frac{3}{10}+3^2\times\frac{2}{5}-4=1$$

$$V(Y)=0^2\times\frac{1}{6}+1^2\times\frac{1}{3}+2^2\times\frac{1}{2}-\left(\frac{4}{3}\right)^2=\frac{5}{9}$$

よって，

$$E(2X-Y)=2E(X)-E(Y)\qquad (X \text{ と } Y \text{ の独立性より})$$

$$=2\times2-\frac{4}{3}=\mathbf{\frac{8}{3}}$$

$$V(2X-Y)=4V(X)+V(Y)=4\times1+\frac{5}{9}=\mathbf{\frac{41}{9}}$$

3.11 (a) 周辺和をとることによって，X と Y の周辺分布は

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

y	-1	0	1	2
$P(Y=y)$	$\frac{2}{8}$	$\frac{2}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

(b) $E(Y)=-1\times\frac{2}{8}+0\times\frac{2}{8}+1\times\frac{3}{8}+2\times\frac{1}{8}=\boldsymbol{\frac{3}{8}}$

$V(Y)=(-1)^2\times\frac{2}{8}+0^2\times\frac{2}{8}+1^2\times\frac{3}{8}+2^2\times\frac{1}{8}-\left(\frac{3}{8}\right)^2=\boldsymbol{\frac{63}{64}}$

(c) $P(X=1,\ Y=-1)=\frac{1}{8}$, $P(X=1)P(Y=-1)=\frac{3}{8}\times\frac{2}{8}$ であるから $P(X=1,\ Y=-1)\neq P(X=1)P(Y=-1)$. よって，$X$, Y は**独立ではない**.

(d) $E(X+Y)=\{1+(-1)\}\times\frac{1}{8}+(1+0)\times\frac{2}{8}+\cdots+(2+2)\times\frac{1}{8}=\boldsymbol{2}$

(e) $E(XY)=1\times(-1)\times\frac{1}{8}+1\times0\times\frac{1}{8}+\cdots+2\times2\times\frac{1}{8}=\boldsymbol{\frac{7}{8}}$

3.12 赤球1個と白球3個がとり出される確率は

$$\frac{{}_2C_1\times{}_5C_3}{{}_8C_4}=\frac{2}{7}$$

袋から4個とり出すとき，x 個の赤球と y 個の青球がとり出される確率，すなわち X と Y の結合確率分布は，

$$P(X=x,\ Y=y)=\frac{{}_2C_x\times{}_1C_y\times{}_5C_{4-x-y}}{{}_8C_4}\qquad(x=0,1,2\,;\ y=0,1)$$

すなわち，

x \ y	0	1	$P(X=x)$
0	$\frac{5}{70}$	$\frac{10}{70}$	$\frac{15}{70}$
1	$\frac{20}{70}$	$\frac{20}{70}$	$\frac{40}{70}$
2	$\frac{10}{70}$	$\frac{5}{70}$	$\frac{15}{70}$
$P(Y=y)$	$\frac{1}{2}$	$\frac{1}{2}$	1

$P(X=1,\ Y=0)\neq P(X=1)P(Y=0)$. よって，$X$ と Y は独立ではない.
$Z=X+Y$ の確率分布は

z	0	1	2	3
$P(Z=z)$	$\frac{5}{70}$	$\frac{30}{70}$	$\frac{30}{70}$	$\frac{5}{70}$

$$E(z)=0\times\frac{5}{70}+1\times\frac{30}{70}+2\times\frac{30}{70}+3\times\frac{5}{70}=\mathbf{\frac{3}{2}}$$

$$V(z)=0^2\times\frac{5}{70}+1^2\times\frac{30}{70}+2^2\times\frac{30}{70}+3^2\times\frac{5}{70}-\left(\frac{3}{2}\right)^2=\mathbf{\frac{15}{28}}$$

3.13 車の販売台数を X, 月間収入を Y とすると, $Y=12+3X$.

$$E(Y)=E(12+3X)=12+3E(X)=12+3\times1.85=\mathbf{17.55}\ (万円)$$

$$V(Y)=V(12+3X)=9V(X)$$

$$\therefore\quad \sqrt{V(Y)}=3\sqrt{V(X)}=3\times1.24=\mathbf{3.72}\ (万円)$$

3.14 (a) $\int_0^1 ax^2dx=\frac{a}{3}=1 \quad\Rightarrow\quad a=\mathbf{3}$

(b) $P\left(X>\frac{1}{2}\right)=\int_{\frac{1}{2}}^1 3x^2dx=3\left[\frac{x^3}{3}\right]_{\frac{1}{2}}^1=\mathbf{\frac{7}{8}}$

(c) $E(X)=\int_0^1 3x^3dx=\mathbf{\frac{3}{4}},\quad V(X)=\int_0^1 3x^4dx-\left(\mathbf{\frac{3}{4}}\right)^2=\mathbf{\frac{3}{80}}$

(d) $F(x)=\begin{cases}\mathbf{0} & (x<0)\\ \boldsymbol{x^3} & (0\leqq x<1)\\ \mathbf{1} & (x\geqq 1)\end{cases}$

3.15 (a) $f(x)$ は密度関数だから, $\int_0^1 f(x)dx=1$. よって

$$1=\int_0^1(a+bx+cx^2)dx=\left[ax+\frac{b}{2}x^2+\frac{c}{3}x^3\right]_0^1=a+\frac{b}{2}+\frac{c}{3} \qquad (1)$$

X の平均が $\frac{1}{2}$ で, 分散が $\frac{1}{20}$ であることから

$$\frac{1}{2}=\int_0^1x(a+bx+cx^2)dx=\left[\frac{a}{2}x^2+\frac{b}{3}x^3+\frac{c}{4}x^4\right]_0^1=\frac{a}{2}+\frac{b}{3}+\frac{c}{4} \qquad (2)$$

$$\frac{1}{20}=\int_0^1x^2(a+bx+cx^2)dx-\left(\frac{1}{2}\right)^2=\left[\frac{a}{3}x^3+\frac{b}{4}x^4+\frac{c}{5}x^5\right]_0^1-\frac{1}{4}=\frac{a}{3}+\frac{b}{4}+\frac{c}{5}-\frac{1}{4} \qquad (3)$$

(1), (2), (3)より

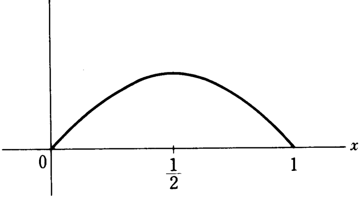
$$\begin{cases}6a+3b+2c=6 & (4)\\ 6a+4b+3c=6 & (5)\\ 20a+15b+12c=18 & (6)\end{cases}$$

これを解けば, $a=\mathbf{0}$, $b=\mathbf{6}$, $c=\mathbf{-6}$.

(b) (a)より, 密度関数は

$$f(x)=6x(1-x) \qquad (0\leqq x\leqq 1)$$

で, グラフは右図に示すように平均 $=\frac{1}{2}$ に関して対称だから, モードもメジアンも平均と同じ $\frac{1}{2}$ である.



3.16 (a) $E(X)=\dfrac{3}{10}\displaystyle\int_1^3 x^2(3-x)dx=\dfrac{3}{10}\left[x^3-\dfrac{x^4}{4}\right]_1^3=\mathbf{1.8}$

$$V(X)=\frac{3}{10}\int_1^3 x^3(3-x)dx-1.8^2=\frac{3}{10}\left[\frac{3}{4}x^4-\frac{x^5}{5}\right]_1^3-1.8^2=\mathbf{0.24}$$

(b) $E(X)=\dfrac{3}{2}\displaystyle\int_0^1 x(1-x^2)dx=\dfrac{3}{2}\left[\dfrac{x^2}{2}-\dfrac{x^4}{4}\right]_0^1=\mathbf{\dfrac{3}{8}}$

$$V(X)=\frac{3}{2}\int_0^1 x^2(1-x^2)dx-\left(\frac{3}{8}\right)^2=\frac{3}{2}\left[\frac{x^3}{3}-\frac{x^5}{5}\right]_0^1-\left(\frac{3}{8}\right)^2=\mathbf{\frac{19}{320}}$$

(c) $E(X)=\displaystyle\int_{-1}^0 x(1+x)dx+\int_0^1 x(1-x)dx$

$$=\left[\frac{x^2}{2}+\frac{x^3}{3}\right]_{-1}^0+\left[\frac{x^2}{2}-\frac{x^3}{3}\right]_0^1=\mathbf{0}$$

$$V(X)=\int_{-1}^0 x^2(1+x)dx+\int_0^1 x^2(1-x)dx$$

$$=\left[\frac{x^3}{3}+\frac{x^4}{4}\right]_{-1}^0+\left[\frac{x^3}{3}-\frac{x^4}{4}\right]_0^1=\mathbf{\frac{1}{6}}$$

3.17 (a) $F(x)=\dfrac{3}{4}\displaystyle\int_{-1}^x(1-t^2)dt=\dfrac{3}{4}\left[t-\dfrac{t^3}{3}\right]_{-1}^x=\dfrac{1}{4}(2+3x-x^3)$

$$\therefore\quad F(x)=\begin{cases}0 & (x<1)\\ \dfrac{1}{4}(2+3x-x^3) & (-1\leqq x<1)\\ 1 & (x\geqq 1)\end{cases}$$

(b) $F(x)=3\displaystyle\int_0^x(1-t)^2dt=3\left[-\dfrac{(1-t)^3}{3}\right]_0^x=1-(1-x)^3$

$$\therefore\quad F(x)=\begin{cases}0 & (x<0)\\ 1-(1-x)^3 & (0\leqq x<1)\\ 1 & (x\geqq 1)\end{cases}$$

(c) $F(x)=\dfrac{3}{32}\displaystyle\int_0^x t(4-t)dt=\dfrac{3}{32}\left[2t^2-\dfrac{t^3}{3}\right]_0^x=\dfrac{1}{32}x^2(6-x)$

$$\therefore\quad F(x)=\begin{cases}0 & (x<0)\\ \dfrac{1}{32}x^2(6-x) & (0\leqq x<4)\\ 1 & (x\geqq 4)\end{cases}$$

3.18 チップ1個の寿命が500時間以内である確率は

$$P(X<500)=\int_{200}^{500}\frac{200}{x^2}dx=200\left[-\frac{1}{x}\right]_{200}^{500}=\frac{3}{5}$$

よって，両方をとり替えねばならない確率は

$$\frac{3}{5}\times\frac{3}{5}=\mathbf{\frac{9}{25}}$$

3.19 (a) $P(0<X<1)=0.03\displaystyle\int_0^1(x^2+3)dx=0.03\left[\dfrac{x^3}{3}+3x\right]_0^1=\mathbf{0.1}$

(b) $P(1<X<2)=0.03\left[\dfrac{x^3}{3}+3x\right]_1^2=0.16$

$$P(2<x<3)=0.03\left[\frac{x^3}{3}+3x\right]_2^3=0.28$$

$$P(3<x<4)=0.03\left[\frac{x^3}{3}+3x\right]_3^4=0.46$$

弾丸を1発射ったときの得点を Y とすると，Y の確率分布は

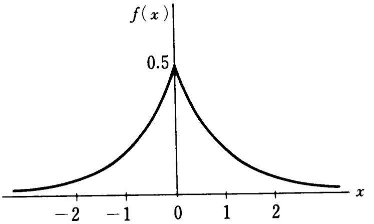
y	0	1	3	5
$P(Y=y)$	0.46	0.28	0.16	0.10

よって，(i) 確率が最も大きいのは **0 点**.

(ii) $E(Y)=0\times0.46+1\times0.28+3\times0.16+5\times0.10=\mathbf{1.26}$ **点**.

3.20 $f(x)$ は

$$f(x)=\begin{cases}\dfrac{1}{2}e^x & (x<0)\\[2mm] \dfrac{1}{2} & (x=0)\\[2mm] \dfrac{1}{2}e^{-x} & (x>0)\end{cases}$$



より，そのグラフは右図のようになる.

$E(X)=\mathbf{0}$ (図から明らか)

$$V(X)=E(X^2)=\frac{1}{2}\int_{-\infty}^0 x^2e^xdx+\frac{1}{2}\int_0^\infty x^2e^{-x}dx$$

$$=\int_0^\infty x^2e^{-x}dx=\Big[-e^{-x}x^2\Big]_0^\infty+2\int_0^\infty xe^{-x}dx=2\int_0^\infty xe^{-x}dx$$

ここで

$$\int_0^\infty xe^{-x}dx=\Big[-e^{-x}x\Big]_0^\infty+\int_0^\infty e^{-x}dx$$

$$=\Big[-e^{-x}\Big]_0^\infty=1$$

となるから

$V(X)=\mathbf{2}$

4 章

4.1 (a) 8回の投げで1の目の出る数を X とすると，X は2項分布

$$P(x)={}_8\mathrm{C}_x\left(\frac{1}{6}\right)^x\left(\frac{5}{6}\right)^{8-x}\qquad(x=0, 1, 2, \cdots, 8)$$

に従う．よって

(i) $P(0)=\left(\dfrac{5}{6}\right)^8=\mathbf{0.233}$

(ii) $P(4)={}_8\mathrm{C}_4\left(\dfrac{1}{6}\right)^4\left(\dfrac{5}{6}\right)^4=\mathbf{0.026}$

(iii) $P\left(X \geqq 2\right) = 1 - P(0) - P(1)$

$$= 1 - 0.233 - 0.372 = \mathbf{0.395}$$

(b) $P(x) = {}_5C_x\left(\frac{1}{3}\right)^x\left(\frac{2}{3}\right)^{5-x} \qquad (x = 0.\ 1, 2, 3, 4, 5)$

ゆえに,

(i) $P(2) = {}_5C_2\left(\frac{1}{3}\right)^2\left(\frac{2}{3}\right)^3 = \mathbf{0.329}$

(ii) $P(X \geqq 4) = P(4) + P(5)$

$$= {}_5C_4\left(\frac{1}{3}\right)^4\frac{2}{3} + {}_5C_5\left(\frac{1}{3}\right)^5 = \mathbf{0.045}$$

(iii) $P(2 \leqq X \leqq 5) = 1 - P(0) - P(1) = \mathbf{0.539}$

4.2 与えられた条件から

$$np = 16, \qquad npq = 3.2$$

よって.

$$q = \frac{3.2}{16} = 0.2$$

したがって,

$$p = 1 - 0.2 = \mathbf{0.8}$$

$$n = \frac{16}{p} = \frac{16}{0.8} = \mathbf{20}$$

2項分布のモードの公式より

$$(n+1)p = (20+1) \times 0.8 = 16.8$$

よって, モードは **16.**

4.3 5人の中の男の子の数を X とすると, X の確率分布は $n=5$, $p=\frac{1}{2}$ の2項分布

$$P(x) = {}_5C_x\left(\frac{1}{2}\right)^x\left(\frac{1}{2}\right)^{5-x}$$

$$= {}_5C_x\left(\frac{1}{2}\right)^5 \qquad (x = 0, 1, 2, \cdots, 5)$$

に従う. よって

(a) $P(X \geqq 4) = P(4) + P(5)$

$$= {}_5C_4\left(\frac{1}{2}\right)^5 + {}_5C_5\left(\frac{1}{2}\right)^5 = \mathbf{\frac{3}{16}}$$

(b) $P(1 \leqq X \leqq 4) = {}_5C_1\left(\frac{1}{2}\right)^5 + {}_5C_2\left(\frac{1}{2}\right)^5 + {}_5C_3\left(\frac{1}{2}\right)^5 + {}_5C_4\left(\frac{1}{2}\right)^5 = \mathbf{\frac{15}{16}}$

(c) $P(X = 0 \text{ または } 5) = P(0) + P(5) = {}_5C_0\left(\frac{1}{2}\right)^5 + {}_5C_5\left(\frac{1}{2}\right)^5 = \mathbf{\frac{1}{16}}$

(d) $P(\mathrm{W, W, M, M, M}) = P(\mathrm{W}) \cdot P(\mathrm{W}) \cdot P(\mathrm{M}) \cdot P(\mathrm{M}) \cdot P(\mathrm{M})$

$$= \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \mathbf{\frac{1}{32}}$$

4.4 1回の試行で正答を得る確率が $p=\dfrac{1}{4}$ のとき，10回の試行で正答を高々2回得る確率を求めればよい．X を10問中得られる正答の数とすると，X は2項分布

$$P(x)={}_{10}\mathrm{C}_x\left(\frac{1}{4}\right)^x\left(\frac{3}{4}\right)^{10-x} \qquad (x=0,1,2,\cdots,10)$$

に従う，よって，

$$\begin{aligned}P(X\leqq 2)&=P(0)+P(1)+P(2)\\&=\left(\frac{3}{4}\right)^{10}+{}_{10}\mathrm{C}_1\left(\frac{1}{4}\right)\left(\frac{3}{4}\right)^9+{}_{10}\mathrm{C}_2\left(\frac{1}{4}\right)^2\left(\frac{3}{4}\right)^8\\&=0.0563+0.1877+0.2816=\mathbf{0.5256}\end{aligned}$$

4.5 1回の投げで6の目の出る確率は $p=\dfrac{1}{6}$ である．n 回の投げで6の目の出る回数を X とすると，X は2項分布

$$P(x)={}_n\mathrm{C}_x\left(\frac{1}{6}\right)^x\left(\frac{5}{6}\right)^{n-x} \qquad (x=0,1,\cdots,n)$$

に従う．題意より，この問題は

$$\sum_{x=1}^{n}{}_n\mathrm{C}_x\left(\frac{1}{6}\right)^x\left(\frac{5}{6}\right)^{n-x}\geqq 0.95$$

を満たす n を求める問題である．この式を整理すれば，

$$1-\sum_{x=1}^{n}{}_n\mathrm{C}_x\left(\frac{1}{6}\right)^x\left(\frac{5}{6}\right)^{n-x}\leqq 0.05$$

$$\left(\frac{5}{6}\right)^n\leqq 0.05$$

両辺の対数をとれば

$$n\geqq\frac{\log 0.05}{\log\dfrac{5}{6}}\fallingdotseq 16.4$$

ゆえに，**17回以上**の投げが必要である．

4.6 X を A が投げた硬貨の表の数，Y を B が投げた硬貨の表の数とすると

$$X\sim B\left(3,\frac{1}{2}\right):\ P(X=x)={}_3\mathrm{C}_x\left(\frac{1}{2}\right)^3 \qquad (x=0,1,2,3)$$

$$Y\sim B\left(4,\frac{1}{2}\right):\ P(X=y)={}_4\mathrm{C}_y\left(\frac{1}{2}\right)^4 \qquad (y=0,1,2,3,4)$$

X と Y は独立で，事象 $X>Y$ は

$$(X=3, Y=2),\ (X=3,\ Y=1), (X=3,\ Y=0)$$
$$(X=2,\ Y=1),\ (X=2,\ Y=0), (X=1,\ Y=0)$$

のときに起こり，これらは互いに排反であるから

$$\begin{aligned}P(X>Y)&=P\{(X=3,\ Y=2)\}+P\{(X=3,\ Y=1)\}+\cdots+P\{(X=1,\ Y=0)\}\\&=P(X=3)P(Y=2)+P(X=3)P(Y=1)+\cdots+P(X=1)P(Y=0)\\&=\frac{1}{8}\times\frac{6}{16}+\frac{1}{8}\times\frac{4}{16}+\frac{1}{8}\times\frac{1}{16}+\frac{3}{8}\times\frac{4}{16}+\frac{3}{8}\times\frac{1}{16}+\frac{3}{8}\times\frac{1}{16}\\&=\frac{29}{128}\fallingdotseq\mathbf{0.227}\end{aligned}$$

4.7 (a) 公式より

$$\mu = np = 8\times\frac{1}{2} = \mathbf{4}$$

$$\sigma = \sqrt{npq} = \sqrt{8\times\frac{1}{2}\times\frac{1}{2}} = \sqrt{2} \fallingdotseq \mathbf{1.414}$$

(b) この2項分布は

$$P(x) = {}_8\mathrm{C}_x\left(\frac{1}{2}\right)^8 \qquad (x=0, 1, 2, \cdots, 8)$$

と書ける.

$$\begin{aligned}P(|X-\mu|\geqq 2\sigma) &= P(|X-4|\geqq 2.828)\\ &= P(X-4\leqq -2.828 \text{ または } X-4\geqq 2.828)\\ &= P(X\leqq 1.172 \text{ または } X\geqq 6.828)\\ &= P(0)+P(1)+P(7)+P(8)\\ &= ({}_8\mathrm{C}_0+{}_8\mathrm{C}_1+{}_8\mathrm{C}_7+{}_8\mathrm{C}_8)\times\left(\frac{1}{2}\right)^8 \fallingdotseq \mathbf{0.07}\end{aligned}$$

(c) $$\begin{aligned}E\{X(X-3)\} &= E(X^2)-3E(X)\\ &= \sigma^2+\mu^2-3\mu\\ &= 2+16-12 = \mathbf{6}\end{aligned}$$

4.8 2項分布

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

の形を調べるために, 隣接する2つの確率の比

$$\frac{P(x+1)}{P(x)} = \frac{{}_n\mathrm{C}_{x+1}p^{x+1}q^{n-x-1}}{{}_n\mathrm{C}_x p^x q^{n-x}} = \frac{(n-x)p}{(x+1)q}$$

を考える. モードは, x が0から n まで動くとき, この比が増加から減少へ変わるときの x の値である. 上の式から

$$(n-x)p \gtreqqless (x+1)q \text{ のとき}, \ P(x+1)\gtreqqless P(x) \qquad (x=0, 1, 2, \cdots, n)$$

上左の不等式は, $q=1-p$ とおいて x について解けば, $x\lesseqqgtr (n+1)p-1)$ となる. よって,

$$x\lesseqqgtr (n+1)p-1 \text{ のとき}, \ P(x+1)\gtreqqless P(x) \qquad (x=0, 1, 2, \cdots, n)$$

ゆえに, $P(x)$ は x から $x+1$ に移るとき, $(n+1)p-1$ より小さい任意の x に対して増加し, $x=(n+1)p-1$ のとき同じ値をとり, $(n+1)p-1$ より大きい任意の x に対して減少する. すなわち, $P(x)$ は最初に増加し, 後で減少する.

以上により, もし $(n+1)p-1$ が整数でないならば, $P(x)$ のモードは $(n+1)p$ を超えない最大の整数であり, $(n+1)p-1$ が整数ならば, $(n+1)p$ も整数で, このとき $P(x)$ は2つのモード $(n+1)p-1$ と $(n+1)p$ をもつ.

4.9 1頁当たり誤字の平均数は $\dfrac{300}{600}=0.5$ であるから, 1頁に含まれる誤字の数 X は, $\lambda=0.5$ のポアソン分布

$$P(x) = \frac{(0.5)^x e^{-0.5}}{x!} \qquad (x=0, 1, 2, \cdots)$$

に従うと考えられる，よって

（a）　$P(2)=\dfrac{0.5^2e^{-0.5}}{2!}=\mathbf{0.0758}$

（b）　$P(X\geqq 2)=1-P(0)-P(1)$

$$=1-e^{-0.5}-0.5e^{-0.5}=\mathbf{0.09}$$

4.10　1日当たりの死者の数を X とすると，題意より X は $\lambda=1.8$ のポアソン分布

$$P(x)=\frac{1.8^x e^{-1.8}}{x!}\qquad (x=0, 1, 2, \cdots)$$

に従うから，

（a）　$P(X\geqq 4)=1-P(X\leqq 3)$

$$=1-P(0)-P(1)-P(2)-P(3)$$

$$=1-e^{-1.8}-1.8e^{-1.8}-\frac{1.8^2}{2!}e^{-1.8}-\frac{1.8^3}{3!}e^{-1.8}$$

$$=1-5.392e^{-1.8}\fallingdotseq\mathbf{0.11}$$

（b）　$P(0)=e^{-1.8}\fallingdotseq\mathbf{0.165}$

4.11　遅刻する学生の数を X とすると，

$$P(x)=\frac{(1.2)^x e^{-1.2}}{x!}\qquad (x=0, 1, 2, \cdots)$$

よって，

（a）　$P(3)=\dfrac{1.2^3e^{-1.2}}{3!}=\mathbf{0.087}$

（b）　$P(X\leqq 1)=P(0)+P(1)$

$$=e^{-1.2}+1.2e^{-1.2}=\mathbf{0.66}$$

4.12　X が従うポアソン分布を

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

とすると，与えられた関係から

$$5=\frac{P(3)}{P(5)}=\frac{\lambda^3e^{-\lambda}}{3!}\div\frac{\lambda^5e^{-\lambda}}{5!}=\frac{20}{\lambda^2}$$

$$\therefore\quad \lambda=2$$

ゆえに，X はポアソン分布

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

に従う，これより

（a）　$P(1)=2e^{-2}=\mathbf{0.271}$

（b）　$P(X\leqq 3)=P(0)+P(1)+P(2)+P(3)$

$$=e^{-2}+2e^{-2}+2e^{-2}+\frac{4}{3}e^{-2}=\mathbf{0.857}$$

4.13　平日の車の需要数を X，週末の車の需要数を Y とし，X，Y はそれぞれ $\lambda=2$, $\lambda=3$ のポアソン分布に従うとする．

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

$$\text{週末}：P(y)=\frac{3^y e^{-y}}{y\,!} \qquad (y=0, 1, 2, \cdots)$$

（a）　P（月曜日に客の申込みを断る）

$=P(X \geqq 6)$

$=1-P(X \leqq 5)$

$=1-e^{-2}-2e^{-2}-2e^{-2}-\frac{4}{3}e^{-2}-\frac{2}{3}e^{-2}-\frac{4}{15}e^{-2}=\mathbf{0.017}$

（b）　P（週末に客の申込みを断る）

$=1-P(Y \leqq 5)$

$=1-e^{-3}-3e^{-3}-\frac{9}{2}e^{-3}-\frac{9}{2}e^{-3}-\frac{27}{8}e^{-3}=\mathbf{0.084}$

4.14　100 個のレンズの中の欠陥品の数を X とすると，$X \sim B(100, 0.015)$.

$n=100$ は十分大きく，$p=0.015$ は非常に小さく，$np=100 \times 0.015=1.5$ は 5 より小さいから，この 2 項分布は $\lambda=np=1.5$ のポアソン分布

$$P(x)=\frac{1.5^x e^{-1.5}}{x\,!} \qquad (x=0, 1, 2, \cdots)$$

で近似できる．よって

（a）　$P(X \leqq 1)=p(0)+p(1)$

$=e^{-1.5}+1.5e^{-1.5}=\mathbf{0.558}$

（b）　$P(X \leqq 4)=1-P(0)-P(1)-P(2)-P(3)$

$=1-e^{-1.5}-1.5e^{-1.5}-1.125e^{+1.5}-0.5625e^{-1.5}=\mathbf{0.066}$

5　章

5.1　（a）　$P(Z<1.64)=\Phi(1.64)=\mathbf{0.9495}$

（b）　$P(Z>1.15)=1-\Phi(1.15)=\mathbf{0.1251}$

（c）　$P(Z<-0.34)=\Phi(-0.34)=1-\Phi(0.34)=\mathbf{0.3669}$

（d）　$P(-1<Z<0.5)=\Phi(0.5)+\Phi(1)-1=\mathbf{0.5328}$

（e）　$P(1.24<Z<2.16)=\Phi(2.16)-\Phi(1.24)=\mathbf{0.0921}$

（f）　$P(Z>-2.19)=\Phi(2.19)=\mathbf{0.98574}$

（g）　$P(Z>c)=0.38 \Rightarrow P(Z<c)=0.62 \Rightarrow c=\mathbf{0.306}$

（h）　$P(Z<c)=0.19 \Rightarrow \Phi(-c)=0.81 \Rightarrow c=\mathbf{-0.878}$

5.2　$X \sim N(3.1) \Rightarrow Z=\frac{X-3}{1} \sim N(0, 1)$

（a）　$P(X<3)=P(Z<0)=\Phi(0)=\mathbf{0.5}$

（b）　$P(X<4.93)=P(Z<1.93)=\Phi(1.93)=\mathbf{0.9732}$

(c) $P(X>3.06)=P(Z<-0.06)=1-\Phi(0.06)=\mathbf{0.4764}$

(d) $P(1<X<4.2)=P(-2<Z<1.2)=\Phi(1.2)+\Phi(2)-1=\mathbf{0.8621}$

(e) $P(1.5<X<2.5)=P(-1.5<Z<-0.5)=\Phi(1.5)-\Phi(0.5)=\mathbf{0.2417}$

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

$$=P(-2<Z<0)=\Phi(2)-\Phi(0)=\mathbf{0.4772}$$

(g) $P(X>c)=0.1 \Rightarrow P(X<c)=0.9 \Rightarrow P\left(Z<\dfrac{c-3}{1}\right)=0.9$

$$\Rightarrow \Phi(c-3)=0.9 \Rightarrow c=\mathbf{4.282}$$

(h) $P(X<c)=0.2 \Rightarrow P\left(Z<\dfrac{c-3}{1}\right)=0.2 \Rightarrow \Phi(c-3)=0.2 \Rightarrow \Phi(3-c)=0.8$

$$\Rightarrow 3-c=0.842 \Rightarrow c=\mathbf{2.158}$$

5.3 (a) $P(X>1530)=P\left(Z>\dfrac{1530-1500}{25}\right)=1-\Phi(1.2)=\mathbf{0.1151}$

(b) $P(X<1480)=P\left(Z<\dfrac{1480-1500}{25}\right)=1-\Phi(0.8)=\mathbf{0.2119}$

(c) $P(1475<X<1550)=P\left(\dfrac{1475-1500}{25}<Z<\dfrac{1550-1500}{25}\right)$

$$=\Phi(2)+\Phi(1)-1=\mathbf{0.8185}$$

5.4 靴下の寿命を X とすると，$X \sim N(55,8^2)$

$$P(X<45)=P\left(Z<\frac{45-55}{8}\right)=1-\Phi(1.25)=0.1056$$

$$P(X<61)=P\left(Z<\frac{61-55}{8}\right)=\Phi(0.75)=0.7734$$

$$5000\times 0.1056=528, \qquad 5000\times 0.7734=3867$$

よって，45 日以内には **528 足**，61 日目以内には **3867 足**.

5.5 部品の長さを X，その平均を μ とすると，$X \sim N(\mu, 2^2)$

(a) $P(X<7.5)=0.975$

$$\Rightarrow P\left(Z<\frac{7.5-\mu}{2}\right)=0.975$$

$$\Rightarrow \Phi\left(\frac{7.5-\mu}{2}\right)=0.975$$

より

$$\frac{7.5-\mu}{2}=1.96 \Rightarrow \mu=\mathbf{3.58}$$

(b) $P(5.4<X<5.5)=P\left(\dfrac{5.4-3.58}{2}<Z<\dfrac{5.5-3.58}{2}\right)$

$$=\Phi(0.96)-\Phi(0.91)=\mathbf{0.0129}$$

5.6 1 箱に詰める品物の重さを X とすると，$X \sim N(52.5, 1.6^2)$

(a) $P(X<50)=P\left(Z<\dfrac{50-52.5}{1.6}\right)\fallingdotseq 1-\Phi(1.56)=\mathbf{0.0594}$

(b) 求める平均を μ とすると，

$$P(X<50)<0.01$$
$$\Rightarrow P(X>50)>0.99$$
$$P\left(Z>\frac{50-\mu}{1.6}\right)>0.99 \Rightarrow \Phi\left(\frac{50-\mu}{1.6}\right)<0.01$$
$$\frac{50-\mu}{1.6}=-2.33 \Rightarrow \mu=\mathbf{53.7}\ (\mathrm{g})$$

5.7 $\bar{x}=\mathbf{34}$(分), $s=\sqrt{40}=\mathbf{6.3}$(分)

（a） 毎日の通学時間を X とすると，$X \sim N(34.40)$

$$P(X>38)=P\left(Z>\frac{38-34}{\sqrt{40}}\right)=1-\Phi(0.63)=\mathbf{0.2643}$$

（b） 求める時間を c とすると，$X \sim N(34, 40)$

$$P(X>c)=0.1$$
$$\Rightarrow\ P\left(Z<\frac{c-34}{\sqrt{40}}\right)=0.9$$
$$\Rightarrow\ \Phi\left(\frac{c-34}{\sqrt{40}}\right)=0.9$$
$$\Rightarrow\ \frac{c-34}{\sqrt{40}}=1.282 \Rightarrow c=\mathbf{42.1}\ (分)$$

5.8 $f(x)$ を 2 回微分すれば

$$f'(x)=-\frac{x-\mu}{\sigma^2}f(x)$$
$$f''(x)=-\frac{1}{\sigma^2}\left\{1-\left(\frac{x-\mu}{\sigma}\right)^2\right\}f(x)$$

$f''(x)=0$ より，$x=\mu\pm\sigma$

$f(x)$ のグラフの凹凸を調べると，次の表のようになる．

x	$<\mu-\sigma$	$\mu-\sigma$	$(\mu-\sigma, \mu+\sigma)$	$\mu+\sigma$	$>\mu+\sigma$
$f''(x)$	$+$	0	$-$	0	$+$
$f(x)$	下に凸	$\frac{1}{\sqrt{2\pi}\sigma}e^{-1}$	上に凸	$\frac{1}{\sqrt{2\pi}\sigma}e^{-1}$	下に凸

よって，$f(x)$ の変曲点の x 座標は $\boldsymbol{\mu\pm\sigma}$.

5.9 おもての出る回数を X とすると，X は 2 項分布

$$P(x)={}_{12}\mathrm{C}_x\left(\frac{1}{2}\right)^{12} \qquad (x=0, 1, 2, \cdots, 12)$$

に従う．よって

（a） $P(X\geqq 9)=P(9)+P(10)+P(11)+P(12)$

$$=({}_{12}\mathrm{C}_9+{}_{12}\mathrm{C}_{10}+{}_{12}\mathrm{C}_{11}+{}_{12}\mathrm{C}_{12})\times\left(\frac{1}{2}\right)^{12}$$
$$=(220+66+12+1)\times\frac{1}{4096}\fallingdotseq\mathbf{0.073}$$

（b） この2項分布の平均と標準偏差は

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

$$\sigma=\sqrt{npq}=\sqrt{12\times\frac{1}{2}\times\frac{1}{2}}\fallingdotseq 1.732$$

$Z=\dfrac{X-6}{1.732}$ とおき，2項分布の正規近似を使えば，

$$P(X\geqq 9)\fallingdotseq P\left(Z>\frac{8.5-6}{1.732}\right)$$
$$=P(Z>1.44)=\mathbf{0.075}$$

5.10 $n=50$ は大きく，$np=50\times0.4=20>5$, $nq=50\times0.6=30>5$ だから，この2項分布は $\mu=20$, $\sigma=\sqrt{50\times0.4\times0.6}=3.46$ の正規分布で近似される．
半整数補正を使って計算する．

（a） $P(X=20)\fallingdotseq P(19.5<Y<20.5)$

$$=P\left(\frac{19.5-20}{3.46}<Z<\frac{20.5-20}{3.46}\right)$$
$$=2\Phi(0.14)-1=\mathbf{0.114}$$

（b） $P(15<X<25)\fallingdotseq P(14.5<Y<25.5)$

$$=P\left(\frac{14.5-20}{3.46}<Z<\frac{25.5-20}{3.46}\right)=2\Phi(1.59)-1=\mathbf{0.8882}$$

5.11 予約した客のうち，当日実際に現れる客の数を X とすると，$X\sim \mathrm{B}\left(50, \dfrac{7}{8}\right)$．よって $\mu=np=50\times\dfrac{7}{8}=43.75$，$\sigma=\sqrt{npq}=\sqrt{50\times\dfrac{7}{8}\times\dfrac{1}{8}}\fallingdotseq 2.34$.

$$P(0\leqq X\leqq 45)\fallingdotseq P\left(\frac{-0.5-43.75}{2.34}<Z<\frac{45.5-43.75}{2.34}\right)$$
$$\fallingdotseq\Phi(0.7479)=\mathbf{0.7736}$$

5.12 $E\{|X|\}=\displaystyle\int_{-\infty}^{\infty}|x|\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{x^2}{2\sigma^2}}dx$

$$=\frac{1}{\sqrt{2\pi}\sigma}\left\{\int_{-\infty}^{0}-xe^{-\frac{x^2}{2\sigma^2}}dx+\int_{0}^{\infty}xe^{-\frac{x^2}{2\sigma^2}}dx\right\}$$
$$=\frac{2}{\sqrt{2\pi}\sigma}\int_{0}^{\infty}xe^{-\frac{x^2}{2\sigma^2}}dx$$

$\dfrac{x}{\sqrt{2}\sigma}=y$ と置けば，$dx=\sqrt{2}\sigma dy$ より

$$E\{|X|\}=\frac{2}{\sqrt{2\pi}\sigma}2\sigma^2\int_{0}^{\infty}ye^{-y^2}dy$$
$$=\frac{4\sigma}{\sqrt{2\pi}}\left[-\frac{1}{2}e^{-y^2}\right]_{0}^{\infty}=\boldsymbol{\sqrt{\frac{2}{\pi}}\sigma}$$

$$V\{|X|\}=E(X^2)-\{E(|X|)\}^2=\sigma^2-\frac{2}{\pi}\sigma^2=\boldsymbol{\left(1-\frac{2}{\pi}\right)\sigma^2}$$

5.13　X は一様分布

$$f(x)=\frac{1}{2} \qquad (-1 \leqq x \leqq 1)$$

に従うから

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

$$=\frac{P\left(\frac{2}{3}<X<1\right)}{2P\left(X>\frac{1}{2}\right)}=\frac{\frac{1}{3}\times\frac{1}{2}}{2\times\frac{1}{2}\times\frac{1}{2}}=\boldsymbol{\frac{1}{3}}$$

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

6　章

6.1　2回の投げの結果を（X_1, X_2）とすると，（X_1, X_2）の標本空間は

(1.1)	(1.2)	(1.3)	(1.4)	(1.5)	(1.6)
(2.1)	(2.2)	(2.3)	(2.4)	(2.5)	(2.6)
(3.1)	(3.2)	(3.3)	(3.4)	(3.5)	(3.6)
(4.1)	(4.2)	(4.3)	(4.4)	(4.5)	(4.6)
(5.1)	(5.2)	(5.3)	(5.4)	(5.5)	(5.6)
(6.1)	(6.2)	(6.3)	(6.4)	(6.5)	(6.6)

であるから，$Y=X_1+X_2$ の標本分布は，

y	2	3	4	5	6	7	8	9	10	11	12
$P(Y=y)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

これより，

$$E(Y)=\mathbf{7} \qquad （分布の対称性より）$$

$$V(Y)=2^2\times\frac{1}{36}+3^2\times\frac{2}{36}+\cdots+12^2\times\frac{1}{36}-7^2=\boldsymbol{\frac{35}{6}}$$

6.2　（a）

X の確率分布：

x	1	2	3
$P(X=x)$	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{6}$

（b）　平均：$\mu=E(X)=1\times\frac{1}{2}+2\times\frac{1}{3}+3\times\frac{1}{6}=\boldsymbol{\frac{5}{3}}$

分散：$\sigma^2=V(X)=1^2\times\frac{1}{2}+2^2\times\frac{1}{3}+3^2\times\frac{1}{6}-\left(\frac{5}{3}\right)^2=\boldsymbol{\frac{5}{9}}$

(c) すべての可能な標本と，それに対する $\bar{x}$ の値と確率を与える表を示す：

標本	$\bar{x}$	確率	標本	$\bar{x}$	確率
(1, 1, 1)	1	$\frac{1}{8}$	(3, 3, 1)	$\frac{7}{3}$	$\frac{1}{72}$
(1, 1, 2)	$\frac{4}{3}$	$\frac{1}{12}$	(3, 1, 3)	$\frac{7}{3}$	$\frac{1}{72}$
(1, 2, 1)	$\frac{4}{3}$	$\frac{1}{12}$	(1, 3, 3)	$\frac{7}{3}$	$\frac{1}{72}$
(2, 1, 1)	$\frac{4}{3}$	$\frac{1}{12}$	(3, 3, 2)	$\frac{8}{3}$	$\frac{1}{108}$
(1, 1, 3)	$\frac{5}{3}$	$\frac{1}{24}$	(3, 2, 3)	$\frac{8}{3}$	$\frac{1}{108}$
(1, 3, 1)	$\frac{5}{3}$	$\frac{1}{24}$	(2, 3, 3)	$\frac{8}{3}$	$\frac{1}{108}$
(3, 1, 1)	$\frac{5}{3}$	$\frac{1}{24}$	(3, 3, 3)	3	$\frac{1}{216}$
(2, 2, 1)	$\frac{5}{3}$	$\frac{1}{18}$	(1, 2, 3)	2	$\frac{1}{36}$
(2, 1, 2)	$\frac{5}{3}$	$\frac{1}{18}$	(1, 3, 2)	2	$\frac{1}{36}$
(1, 2, 2)	$\frac{5}{3}$	$\frac{1}{18}$	(2, 1, 3)	2	$\frac{1}{36}$
(2, 2, 2)	2	$\frac{1}{27}$	(2, 3, 1)	2	$\frac{1}{36}$
(2, 2, 3)	$\frac{7}{3}$	$\frac{1}{54}$	(3, 1, 2)	2	$\frac{1}{36}$
(2, 3, 2)	$\frac{7}{3}$	$\frac{1}{54}$	(3, 2, 1)	2	$\frac{1}{36}$
(3, 2, 2)	$\frac{7}{3}$	$\frac{1}{54}$			

$\overline{X}$の標本分布は

$\bar{x}$	1	$\frac{4}{3}$	$\frac{5}{3}$	2	$\frac{7}{3}$	$\frac{8}{3}$	3
$P(\overline{X}=\bar{x})$	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{7}{24}$	$\frac{11}{54}$	$\frac{7}{72}$	$\frac{1}{36}$	$\frac{1}{216}$

これより

$$E(\overline{X})=1\times\frac{1}{3}+\frac{4}{3}\times\frac{1}{4}+\cdots+3\times\frac{1}{216}=\mathbf{\frac{5}{3}}\ (=\mu)$$

$$V(\overline{X})=1^2\times\frac{1}{8}+\left(\frac{4}{3}\right)^2\times\frac{1}{4}+\cdots+3^2\times\frac{1}{216}-\left(\frac{5}{3}\right)^2=\mathbf{\frac{5}{27}}\ \left(=\frac{\sigma^2}{3}\right)$$

6.3 (a) $E(X-Y)=E(X)-E(Y)=5-7=-2$

$V(X-Y)=V(X)+V(Y)=9+16=25$

$\therefore\ X-Y \sim \boldsymbol{N(-2, 25)}$

(b) $E(3X+Y)=3E(X)+E(Y)=3\times5+7=22$

$V(3X+Y)=9V(X)+V(Y)=9\times9+16=97$

$\therefore\ 3X+Y \sim \boldsymbol{N(22, 97)}$

(c) $E(\overline{X})=E(X)=5,\quad V(\overline{X})=\dfrac{\sigma^2}{n}=\dfrac{9}{25}=0.36\quad \therefore\ \overline{X} \sim \boldsymbol{N(5, 0.36)}$

(d) $E(\overline{X}-\overline{Y})=E(\overline{X})-E(\overline{Y})=E(X)-E(Y)=-2$

$V(\overline{X}-\overline{Y})=V(\overline{X})+V(\overline{Y})=\dfrac{9}{25}+\dfrac{16}{16}=1.36$

$\therefore\ \overline{X}-\overline{Y} \sim \boldsymbol{N(-2, 1.36)}$

6.4 (a) 2人の身長を X_1, X_2 とすると，仮定より

$$X_i \sim N(163.0, 8^2)\qquad (i=1, 2)$$

したがって，

$$X_1-X_2 \sim N(0, 8^2+8^2)$$

ゆえに，

$$\begin{aligned}P(|X_1-X_2|>5)&=2P(X_1-X_2>5)\\&=2P\left(Z>\frac{5-0}{\sqrt{128}}\right)\\&=2P(Z>0.44)\fallingdotseq\mathbf{0.66}\end{aligned}$$

(b) (a)と同様に，2人の身長を Y_1, Y_2 とすると，

$$Y_1-Y_2 \sim N(0, 6^2+6^2)$$

よって，

$$\begin{aligned}P(|Y_1-Y_2|>5)&=2P(Y_1-Y_2>5)\\&=2P\left(Z>\frac{5-0}{\sqrt{72}}\right)\\&=2P(Z>0.59)\fallingdotseq\mathbf{0.56}\end{aligned}$$

(c) 男子生徒の身長を X，女子生徒の身長を Y とすると

$$X \sim N(163.0, 8^2),\qquad Y \sim N(155.5, 6^2)$$

よって，

$$X-Y \sim N(7.5, 8^2+6^2)$$

ゆえに，

$$\begin{aligned}P(|X-Y|>10)&=2P(X-Y>10)\\&=2P\left(Z>\frac{10-7.5}{\sqrt{100}}\right)\\&=2P(Z>0.25)\fallingdotseq\mathbf{0.80}\end{aligned}$$

6.5 男子の身長を X, 女子の身長を Y とすると, $X \sim N(169.3, 4.6^2), Y \sim N(156.6, 4.2^2)$.

(a) $E(\overline{X})=E(X)=\mathbf{169.3}$ (cm) , $\sqrt{V(\overline{X})}=\dfrac{4.6}{\sqrt{64}}=\mathbf{0.575}$ (cm)

$E(\overline{Y})=E(Y)=\mathbf{156.6}$ (cm) , $\sqrt{V(\overline{Y})}=\dfrac{4.2}{\sqrt{36}}=\mathbf{0.7}$ (cm)

(b) $E(\overline{X}-\overline{Y})=E(\overline{X})-E(\overline{Y})=169.3-156.6=\mathbf{12.7}$ (cm)

$\sqrt{V(\overline{X}-\overline{Y})}=\sqrt{V(\overline{X})+V(\overline{Y})}=\sqrt{\dfrac{4.6^2}{64}+\dfrac{4.2^2}{36}}=\mathbf{0.906}$ (cm)

(c) $\overline{X}-\overline{Y} \sim N(12.7, 0.906^2)$

$P(\overline{X}-\overline{Y}>12)=P\left(Z>\dfrac{12.7-12}{0.906}\right)=\Phi(0.773)=\mathbf{0.7803}$

6.6 桃缶の中味の重さを X, 缶の重さを Y とすると

$$X \sim N(300, 1.2^2), \qquad Y \sim N(20, 0.5^2)$$

桃缶1個の重さを W とすると, $W=X+Y$. よって,

(a) $E(W)=E(X+Y)$

$=E(X)+E(Y)$

$=300+20=\mathbf{320}$ (g)

$\sqrt{V(W)}=\sqrt{V(X+Y)}=\sqrt{1.2^2+0.5^2}=\mathbf{1.3}$ (g)

(b) $W \sim N(320, 1.3^2)$

$P(W>323)=P\left(Z>\dfrac{323-320}{1.3}\right)=1-\Phi(2.31)=\mathbf{0.01104}$

6.7 ボルトの直径を X, 穴の直径を Y とすると,

$$X \sim N(2.80, 0.03^2), \qquad Y \sim N(2.90, 0.04^2)$$

ボルトが穴にはまらないのは, $X>Y$ すなわち $X-Y>0$ のときである. $X-Y \sim N(-0.10, 0.03^2+0.04^2)$ より

(a) $P(X-Y>0)=P\left(Z>\dfrac{0-(-0.10)}{0.05}\right)=1-\Phi(2)=\mathbf{0.0228}$

(b) $(1-0.0228)^5=\mathbf{0.89}$

6.8 $X_1, X_2, \cdots, X_n$ は $N(\mu, \sigma^2)$ からの無作為標本であるから, $E(X_i)=\mu$, $V(X_i)=\sigma^2$ $(i=1, 2, \cdots, n)$.

よって,

$$\begin{aligned}E(Y)&=E(X_1+2X_2+\cdots+nX_n)\\&=E(X_1)+2E(X_2)+\cdots+nE(X_n)\\&=(1+2+\cdots+n)\mu\\&=\frac{n(n+1)}{2}\mu\end{aligned}$$

$$\begin{aligned}V(Y)&=V(X_1+2X_2+\cdots+nX_n)\\&=V(X_1)+2^2V(X_2)+\cdots+n^2V(X_n) \qquad (X_1, X_2, \cdots, X_n \text{ は独立だから})\end{aligned}$$

$$=(1^2+2^2+\cdots+n^2)\sigma^2$$
$$=\frac{n(n+1)(2n+1)}{6}\sigma^2$$

Y は正規変数の 1 次結合であるから，正規分布に従う．

よって，

$$Y \sim \boldsymbol{N}\left(\frac{\boldsymbol{n(n+1)}}{\boldsymbol{2}}\boldsymbol{\mu},\ \frac{\boldsymbol{n(n+1)(2n+1)}}{\boldsymbol{6}}\boldsymbol{\sigma^2}\right)$$

6.9 (a) $\overline{X}$ の標本分布は $N\left(\mu, \dfrac{\sigma^2}{4}\right)$ であるから

$$P(|\overline{X}-\mu|<\sigma)=P(\mu-\sigma<\overline{X}<\mu+\sigma)$$
$$=P\left(\frac{\mu-\sigma-\mu}{\frac{\sigma}{2}}<Z<\frac{\mu+\sigma-\mu}{\frac{\sigma}{2}}\right)$$
$$=2\Phi(2)-1=\mathbf{0.9545}$$

(b) $n=64$ は十分大きいから，中心極限定理により，$\overline{X} \sim N\left(80, \dfrac{100}{64}\right)$.

$$P(\overline{X}>82)=P\left(Z>\frac{82-80}{\frac{10}{8}}\right)=1-\Phi(1.6)=\mathbf{0.0548}$$

6.10 (a) まず，与えられた母集団分布の平均と分散を求める．

$$\mu=6\int_0^1 x^2(1-x)\,dx=6\left[\frac{x^3}{3}-\frac{x^4}{4}\right]_0^1=\frac{1}{2}$$
$$\sigma^2=6\int_0^1 x^3(1-x)\,dx-\left(\frac{1}{2}\right)^2=6\left[\frac{x^4}{4}-\frac{x^5}{5}\right]_0^1-\frac{1}{4}=\frac{1}{20}$$

次に，$\overline{X}$ の平均と分散を求める．

$$E(\overline{X})=\mu=\frac{1}{2},\qquad V(\overline{X})=\frac{\sigma^2}{n}=\frac{\frac{1}{20}}{45}=\frac{1}{900}$$

$n=45$ は大きいから，中心極限定理によって

$$\overline{X} \sim \boldsymbol{N}\left(\frac{\mathbf{1}}{\mathbf{2}}, \frac{\mathbf{1}}{\mathbf{900}}\right)$$

したがって，

(b) $P(\bar{X}<0.55)=P\left(Z<\dfrac{0.55-0.5}{1/30}\right)=P(Z<1.5)=\mathbf{0.4332}$

(c) $P(\bar{X}>0.46)=P\left(Z>\dfrac{0.46-0.5}{1/30}\right)=P(Z>-1.2)=\mathbf{0.3849}$

6.11 一様分布の平均と分散の公式から，$\mu=0$，$\sigma^2=\dfrac{(b-a)^2}{12}=\dfrac{\left(\frac{1}{2}+\frac{1}{2}\right)^2}{12}=\dfrac{1}{12}$.

よって $E(\overline{X})=\mu=0$，$V(\overline{X})=\dfrac{\sigma^2}{n}=\dfrac{1}{12n}$.

中心極限定理より，$\overline{X} \sim N\left(0, \dfrac{1}{12n}\right)$.

$$\therefore \quad P\left(|\overline{X}|<\frac{1}{2\sqrt{n}}\right)=P\left(-\frac{1}{2\sqrt{n}}<\overline{X}<\frac{1}{2\sqrt{n}}\right)$$

$$=P\left(\frac{-\frac{1}{2\sqrt{n}}-0}{\frac{1}{\sqrt{12n}}}<Z<\frac{\frac{1}{2\sqrt{n}}-0}{\frac{1}{\sqrt{12n}}}\right)=2\Phi(\sqrt{3})-1=\mathbf{0.9168}$$

7 章

7.1 (a) 平均と分散の不偏推定値は $\bar{x}$ と u^2 であるから,

$$\bar{x}=\frac{1}{200}\{4.5\times3+14.5\times11+\cdots+94.5\times5\}=\mathbf{50.12}$$

$$u^2=\frac{1}{200-1}\left\{\frac{4.5^2\times3+\cdots+94.5^2\times5}{200}-50.12^2\right\}=388.09$$

$$\therefore \quad u=\mathbf{19.70}$$

(b) $n=200$ は大標本であるから, μ の 95 % 信頼区間は

$$\bar{x}-1.96\sqrt{\frac{u^2}{n}}<\mu<\bar{x}+1.96\sqrt{\frac{u^2}{n}}$$

$$50.12-1.96\sqrt{\frac{388.09}{200}}<\mu<50.12+1.96\sqrt{\frac{388.09}{200}}$$

よって, 平均 μ の 95%の信頼限界は

$$\mathbf{47.39,\ 52.85}$$

(c)

$$1.96\sqrt{\frac{u^2}{n}}=1.96\times\frac{19.70}{\sqrt{n}}=2$$

より

$$n=370.6$$

よって 最小の n は **371**.

7.2 $\overline{X}$ は μ の不偏推定量であるが, $\overline{X}^2$ は μ^2 の不偏推定量ではない. なぜなら,

$$E(\overline{X}^2)=V(\overline{X})+\{E(\overline{X})\}^2=\frac{\sigma^2}{n}+\mu^2=\frac{1}{n}+\mu^2\neq\mu^2$$

上の式より, $E\left(\overline{X}^2-\frac{1}{n}\right)=\mu^2$ であるから, $\overline{X}^2-\frac{1}{n}$ が μ^2 の不偏推定量である.

7.3 (a) X の平均 μ は

$$\mu=E(X)=\int_0^\theta x\cdot\frac{2x}{\theta^2}dx=\frac{2}{\theta^2}\int_0^\theta x^2\,dx=\frac{2}{3}\theta$$

で, $E(\overline{X})=\mu=\frac{2}{3}\theta$ であるから, $E\left(\frac{3}{2}\overline{X}\right)=\theta$.

ゆえに, θ の不偏推定量は $\frac{3}{2}\overline{X}$ で, その不偏推定値は

$$\frac{3}{2}\times\frac{1}{5}(0.6+1.5+0.8+1.1+1.3)=\frac{3}{2}\times1.06=\mathbf{1.59}$$

(b) $E(X^2)=\int_0^{\theta} x^2\cdot\frac{2x}{\theta^2}\,dx=\frac{2}{\theta^2}\int_0^{\theta} x^3\,dx=\frac{1}{2}\theta^2$

$$\sigma^2\equiv V(X)=E(X^2)-\mu^2=\frac{1}{2}\theta^2-\left(\frac{2}{3}\theta\right)^2=\frac{1}{18}\theta^2$$

よって，σ^2 の標本不偏分散を U^2 とすると，

$$E(U^2)=\sigma^2=\frac{1}{18}\theta^2$$

したがって，$18U^2$ が θ^2 の不偏推定量である．

$$u^2=\frac{1}{n-1}\{\sum x_i^2-n\bar{x}^2\}$$

$$=\frac{1}{3}\{(0.6^2+1.5^2+\cdots+1.3^2)-5\times1.06^2\}=0.177$$

よって，θ^2 の不偏推定値は

$$18\times0.177=\mathbf{3.186}$$

7.4 尤度関数 $L(\theta)$ は，

$$L(\theta)=\prod_{i=1}^{n} f(X_i\,;\theta)=(1+\theta)^n(X_1X_2\cdots X_n)^{\theta}$$

両辺の対数をとれば，

$$\log L(\theta)=n\log(1+\theta)+\theta\sum_{i=1}^{n}\log X_i$$

$$\frac{d\log(\theta)}{d\theta}=\frac{n}{1+\theta}+\sum_{i=1}^{n}\log X_i=0$$

$$\therefore\quad \hat{\theta}=-\left(\frac{n}{\sum_{i=1}^{n}\log X_i}+1\right)$$

7.5 度数分布表より，簡便計算法によって平均と分散を求める．

x	f	u	uf	u^2f
17.75	2	−5	−10	50
18.25	5	−4	−20	80
18.75	8	−3	−24	72
19.25	13	−2	−26	52
19.75	19	−1	−19	19
$x_0=$20.25	25	0	0	0
20.75	15	1	15	15
21.25	7	2	14	28
21.75	5	3	15	45
22.25	0	4	0	0
22.75	1	5	5	25
計	100		−50	386

$n=100,\ x_0=20.25,\ c=0.5$ だから

$$\bar{x}=c\frac{\sum u_i f_i}{n}+x_0=0.5\times\frac{-50}{100}+20.25=20.00$$

$$u=c\sqrt{\frac{1}{n-1}\{\sum u_i^2 f_i-\frac{1}{n}(\sum u_i f_i)^2\}}=0.5\sqrt{\frac{1}{99}\left\{386-\frac{(-50)^2}{100}\right\}}=0.95$$

ゆえに，95%信頼限界は，

$$\bar{x}\pm z_{0.025}\frac{u}{\sqrt{n}}=20.00\pm1.96\times\frac{0.95}{\sqrt{100}}=\mathbf{19.8,\ 20.2}$$

7.6 （a） μ の不偏推定量は標本平均 $\overline{X}$ で，σ^2 の不偏推定量は標本不偏分散 U^2 だから，

$$\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i=\frac{1}{12}\times72=6$$

$$u^2=\frac{1}{n-1}\left\{\sum x_i{}^2-\frac{1}{n}(\sum x_i)^2\right\}=\frac{1}{12-1}\left\{1620-\frac{1}{12}\times72^2\right\}=108$$

よって，μ と σ^2 の不偏推定値は，それぞれ **6** と **108** である．

（b） μ の 95%の信頼区間は，

$$\bar{x}-t_{0.025}(n-1)\frac{u}{\sqrt{n}}<\mu<\bar{x}+t_{0.025}(n-1)\frac{u}{\sqrt{n}}$$

$n=12$,　$u=12-1=11$, t 分布表より　$t_{0.025}\,(11)=2.201$.

この値と $\bar{x}=6,\ u^2=108$ を代入すれば

$$6-2.201\sqrt{\frac{108.0}{12}}<\mu<6+2.201\sqrt{\frac{108.0}{12}}$$

$$\mathbf{-0.603<\mu<12.603}$$

次に，σ^2 の 95%信頼区間は，

$$\frac{(n-1)u^2}{\chi^2_{0.025}(n-1)}<\sigma^2<\frac{(n-1)u^2}{\chi^2_{0.975}(n-1)}$$

$n=12$, χ^2 分布表より $\chi^2_{0.025}(11)=21.92,\ \chi^2_{0.975}(11)=3.82,\ u^2=108.0$

よって，

$$\frac{11\times108.0}{21.92}<\sigma^2<\frac{11\times108.0}{3.82}$$

$$\mathbf{54.2<\sigma^2<311.0}$$

7.7 AとBの真の重さを，それぞれ μ_1, μ_2 とすると，分散が既知の場合の $\mu_1-\mu_2$ に対する 95％信頼区間は，

$$\bar{x}_1-\bar{x}_2-1.96\sqrt{\frac{\sigma_1{}^2}{n_1}+\frac{\sigma_2{}^2}{n_2}}<\mu_1-\mu_2<\bar{x}_1-\bar{x}_2+1.96\sqrt{\frac{\sigma_1{}^2}{n_1}+\frac{\sigma_2{}^2}{n_2}}$$

与えられた情報より，

$$n_1=10,\quad \bar{x}_1=8.6,\quad \sigma_1^2=0.5^2$$
$$n_2=15,\quad \bar{x}_2=6.8,\quad \sigma_2^2=0.5^2$$

であるから，

$$8.6-6.8-1.96\sqrt{\frac{0.5^2}{10}+\frac{0.5^2}{15}}<\mu_1-\mu_2<8.6-6.8+19.6\sqrt{\frac{0.5^2}{10}+\frac{0.5^2}{15}}$$

$$\mathbf{1.4<\mu_1-\mu_2<2.2}$$

7.8　容器こみ製品１個の重さを x とすると，$x=x_1+x_2$ であるから，x は，平均 $\mu_1+\mu_2=30.0+5.0=35.0$, 標準偏差 $\sqrt{\sigma_1{}^2+\sigma_2{}^2}=\sqrt{0.4^2+0.3^2}=0.5$ の正規分布に従う．

よって，x の平均 μ_x の 95％信頼区間は，

$$35-1.96\times 0.5<\mu_x<35+1.96\times 0.5$$

$$\mathbf{34.02<\mu_x<35.98}$$

製品 10 個入り１箱の重さを y とする，$y=10x+x_3$ は，平均 $10\times 35+35=385$, 標準偏差 $\sqrt{10\times 0.25+1.0^2}=1.87$ の正規分布に従うから，y の平均 μ_y の 95％信頼区間は，

$$385-1.96\times 1.87<\mu_y<385+1.96\times 1.87$$

$$\mathbf{381.3<\mu_y<388.7}$$

7.9　この学校で，めがねをかけている生徒の割合を p とすると，

$$\text{標本比率 } \hat{p}=\frac{14}{40}=0.35$$

（a）　p の 95％信頼区間は，

$$0.35-1.96\sqrt{\frac{0.35\times 0.65}{40}}<p<0.35+1.96\sqrt{\frac{0.35\times 0.65}{40}}$$

$$\mathbf{0.202<p<0.498}$$

（b）　めがねをかけている生徒の総数を M とすると

$$M=1800p$$

（a）で求めた p の信頼区間に 1800 をかければ M に対する 95％信頼区間は

$$\mathbf{364<M<896}$$

7.10　与えられた情報より　$\hat{p}_1-\hat{p}_2=\dfrac{750}{1000}-\dfrac{520}{800}=0.75-0.65=0.10$

$$\sqrt{V(\hat{p}_1-\hat{p}_2)}\fallingdotseq\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1}+\frac{\hat{p}_2\hat{q}_2}{n_2}}=\sqrt{\frac{0.75\times 0.25}{1000}+\frac{0.65\times 0.35}{800}}=0.0217$$

よって，$\hat{p}_1-\hat{p}_2$ の 95％信頼限界は

$$\hat{p}_1-\hat{p}_2\pm 1.96\sqrt{V(\hat{p}_1-\hat{p}_2)}=0.10\pm 1.96\times 0.0217=0.057,\ 0.143$$

ゆえに，95%信頼区間は（**0.057，0.143**）．

7.11　尤度関数 $L(\theta)$ は，

$$L(\theta)=\prod_{i=1}^{5}\theta e^{-\theta x_i}=\theta^5 e^{-\theta(\Sigma x_i)}$$

$\log L(\theta)=5\log\theta-\theta\sum_{i=1}^{5}x_i$ より $\dfrac{d\log L(\theta)}{d\theta}=0$ を解いて，

$$\hat{\theta}=\frac{5}{\sum_{i=1}^{5}x_i}=\frac{1}{\bar{x}}$$

よって，θ の最大推定量は $\dfrac{1}{\bar{x}}$ で，その推定値は $\bar{x}=7.96$ より

$$\hat{\theta}=\frac{1}{7.96}\fallingdotseq\mathbf{0.13}$$

8 章

8.1 (a) $p=\frac{x}{n}=\frac{90}{120}=0.75,\quad z=\frac{\hat{p}-p}{\sqrt{\frac{p(1-p)}{n}}}=\frac{0.75-0.7}{\sqrt{\frac{0.7\times 0.3}{120}}}=1.20$

両側検定の$\sigma=0.05$の棄却域は$|z|>1.96$で，$z=1.20$はこれに入らないからH_0は**採択**.

(b) $\hat{p}_1=\frac{x_1}{n_1}=\frac{118}{420}=0.281,\quad \hat{p}_2=\frac{x_2}{n_2}=\frac{121}{530}=0.228$

$$\hat{p}=\frac{x_1+x_2}{n_1+n_2}=\frac{239}{950}=0.252$$

$$z=\frac{\hat{p}_1-\hat{p}_2}{\sqrt{p(1-p)\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}=\frac{0.281-0.228}{\sqrt{0.252\times 0.748\left(\frac{1}{420}+\frac{1}{530}\right)}}=1.87$$

片側検定の棄却域は$z>1.645$で，1.87はこれに入るからH_0は**棄却**.

8.2 題意より，両側検定が適当である

$$H_0:\ \mu=15$$
$$H_1:\ \mu\neq 15$$
$$u^2=\frac{n}{n-1}s^2=\frac{12}{11}\times 2.1^2=4.81$$
$$\therefore\quad u=2.19$$
$$t=\frac{\bar{x}-\mu}{\frac{u}{\sqrt{n}}}=\frac{17-15}{\frac{2.19}{\sqrt{12}}}=3.16$$

$\alpha=0.05$, $\nu=n-1=12-1=11$. t分布表より，$t_{0.025}(11)=2.201$であるから，この検定の棄却域は$|t|>2.201$. $t=3.16$は棄却域に落ちるからH_0は棄却される．したがって，機械は**順調に作動していない**.

8.3 この機械が切りとる部品の長さをXとし，Xは正規分布に従うとする．機械は正しく調整されているという仮説を立てて，検定は

$$H_0:\ \mu=5.00$$
$$H_1:\ \mu\neq 5.00$$

の両側検定で行う．

与えられたデータより

$$\sum x_i=34.82,\quad \sum x_i^2=173.2102$$
$$u=\sqrt{\frac{1}{6}\left(173.2102-\frac{34.82^2}{7}\right)}\fallingdotseq 0.0306$$

よって，検定統計量の値は

$$t=\frac{\bar{x}-\mu_0}{\frac{u}{\sqrt{n}}}=\sqrt{7}\times\frac{4.97-5.00}{0.0306}\fallingdotseq -2.59$$

両側検定で，$\alpha=0.05$, $\nu=7-1=6$より，$t_{0.025}(6)=2.447$であるから，棄却域は

$$|t| > 2.447$$

$t=2.59$ はこの棄却域に入るから，仮説は棄却される．
機械は**正しく調整されていない**と判断される．

8.4　タイヤの寿命 X は正規分布に従うとすると，この問題は X の平均 μ に関する片側検定の問題である．題意より

$$H_0:\ \mu=20000$$
$$H_1:\ \mu>20000$$

データから，$n=100$, $\bar{x}=21000$, $s=1500$ である．
大標本の場合であるから，$u \fallingdotseq s$．検定統計量は，

$$z=\frac{\bar{x}-\mu_0}{\dfrac{s}{\sqrt{n}}}=\frac{21000-20000}{\dfrac{1500}{\sqrt{100}}}=6.67$$

$\alpha=0.05$ に対する片側検定の棄却域は $z>1.645$．
したがって H_0 は棄却されるから，新型タイヤの寿命は 20,000km を**超える**といえる．

8.5　与えられた情報より，$n_A=100$, $\bar{x}_A=173.2$, $s_A=1.1$
$n_B=100$, $\bar{x}_B=174.1$, $s_B=1.4$
A, B それぞれの機械が詰める粉乳の量を X_A, X_B で表し，$X_A \sim N(\mu_A, \sigma_A{}^2)$, $X_B \sim N(\mu_B, \sigma_B{}^2)$ とする．
題意より，この問題は

$$H_0:\ \mu_A=\mu_B$$
$$H_1:\ \mu_A \neq \mu_B$$

の両側検定で処理するのが適当である．検定統計量の値は，

$$z=\frac{\bar{x}_A-\bar{x}_B}{\sqrt{\dfrac{s_A{}^2}{n_A}+\dfrac{s_B{}^2}{n_B}}}=\frac{173.2-174.1}{\sqrt{\dfrac{1.1^2}{100}+\dfrac{1.4^2}{100}}}=-5.05$$

$$\alpha=0.01, \qquad z_{0.005}=2.576$$

ゆえに，棄却域は $|z| \geqq 2.576$ だから H_0 は棄却される．
よって，2台の機械が詰める粉乳の量には**有意な差がある**．

8.6　(a)　対応する観測値を d_i $(i=1, 2, \cdots, 8)$ とする．

$$\sum d_i=4+3+\cdots+(-3)=18, \qquad \bar{d}=\frac{18}{8}=2.25$$

$$\sum d_i{}^2=4^2+3^2+\cdots+(-3)^2=116$$

$$u_d=\sqrt{\frac{1}{7}\left(116-\frac{18^2}{8}\right)}=3.28$$

これから，

$$t=\frac{\bar{d}-0}{\dfrac{u_d}{\sqrt{n}}}=\frac{2.25}{\dfrac{3.28}{\sqrt{8}}}=1.94$$

$\alpha=0.05$, $\nu=7$, $t_{0.025}(7)=2.365$. $1.94<t_{0.025}(7)$ であるから 2 つの飼育法による味の良さに**有意な差はない**.

（b） **正規分布**の仮定が必要.

8.7 母分散を σ^2 とすると，帰無仮説は H_0：$\sigma^2=10.0$ で，対立仮説は H_1：$\sigma^2\neq 10.0$ の両側検定となる. H_0 の下では，統計量 $\dfrac{ns^2}{\sigma^2}$ は自由度 $n-1$ の χ^2 分布に従う. 有意水準 5% では，χ^2 分布表から，片裾 2.5% ずつの点は 8.91 と 32.85 であるから，この検定の棄却域は $\chi^2>32.85$ または $\chi^2<8.91$ となる.

$$\chi^2=\frac{ns^2}{\sigma^2}=\frac{20\times 5.8}{10}=11.6$$

は棄却域に入らないから，H_0 は採択される. よって分散は **10.0 とみなせる**.

8.8 （a）
$$F=\frac{\max(u_1^2,\ u_2^2)}{\min(u_1^2,\ u_2^2)}=\frac{7.5^2}{4.2^2}=3.19$$

両側検定だから，F 分布の 2.5% の表より，$F_{0.025}(8, 11)=2.95$, $F=3.19$ は棄却域 $F>2.95$ に入っているから，H_0 は**棄却される**.

（b） この場合は片側検定であるから，F 分布の 5% の表より，$F_{0.05}(8, 11)=3.66$. $F=3.19$ は棄却域 $F>3.66$ に入らないから，H_0 は**棄却できない**.

8.9 題意より，比率 p に関する左片側検定

$$H_0:\ p=0.40$$
$$H_1:\ p<0.40$$

の問題である. $n=105$, $x=36$ より，

$$\hat{p}=\frac{36}{105}=0.343$$

$$z=\frac{\hat{p}-p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}}=\frac{0.343-0.40}{\sqrt{\dfrac{0.4\times 0.6}{105}}}=-1.19$$

左片側検定で，$z_{0.05}=1.645$ であるから，棄却域は $z<-1.645$, -1.19 は棄却域に入らないから，H_0 は棄却されない. したがって，学力が**劣っているとはいえない**.

8.10 2 つの比率の差の検定の問題で，題意より両側検定で処理するのが適当である.

$$H_0:\ p_1=p_2$$
$$H_1:\ p_1\neq p_2$$

$n_1=200$, $x_1=48$, $n_2=300$, $x_2=105$ より，

$$\hat{p}_1=\frac{48}{200}=0.24,\quad \hat{p}_2=\frac{105}{300}=0.35,\quad \hat{p}=\frac{x_1+x_2}{n_1+n_2}=\frac{48+105}{200+300}=0.306$$

よって，

$$z=\frac{\hat{p}_1-\hat{p}_2}{\sqrt{\hat{p}\left(\dfrac{1}{n_1}+\dfrac{1}{n_2}\right)}}=\frac{0.24-0.35}{\sqrt{0.306\left(\dfrac{1}{200}+\dfrac{1}{300}\right)}}=-2.18$$

$z_{0.025}=1.96$ で，棄却域は $|z|>1.96$. -2.18 は棄却域に入るから，H_0 は棄却される.

したがって，この番組の視聴率は男と女で**異なる**.

8.11 X は2項分布

$$p(x)=\frac{3!}{x!(3-x)!}p^x(1-p)^{3-x} \qquad (x=0,1,2,3)$$

に従う．題意より

$$H_0:\ p=0.5$$
$$H_1:\ p=0.8$$

(a) 棄却域が $x=3$ のとき，

$$\alpha=P(x=3\mid p=0.5)=\frac{3!}{3!0!}(0.5)^3(1-0.5)^0=\mathbf{0.125}$$

$$\beta=P(x\neq 3\mid p=0.7)=1-P(x=3\mid p=0.7)$$
$$=1-\frac{3!}{3!0!}(0.7)^3(1-0.3)^0=1-0.343=\mathbf{0.657}$$

(b) 棄却域が $x=0$ のとき

$$\alpha=P(x=0\mid p=0.5)=\frac{3!}{0!3!}(0.5)^0(1-0.5)^3=\mathbf{0.125}$$

$$\beta=P(x\neq 0\mid p=0.7)=1-P(x=0\mid p=0.7)$$
$$=1-\frac{3!}{0!3!}(0.7)^0(0.3)^3=1-0.027=\mathbf{0.973}$$

8.12 とり出した5個の中の黒球の数を X とする．X の数で H_0 の棄却，採択をするから，第1種，第2種の過誤の確率をそれぞれ α,β とすれば

$$\alpha=P(H_0\text{ を棄却}\mid H_0\text{ が真})=P(X\neq 5\mid \text{白球 20 個, 黒球 80 個})$$
$$=1-P(X=5\mid \text{白球 20 個, 黒球 80 個})$$
$$=1-\frac{80}{100}\times\frac{79}{99}\times\frac{78}{98}\times\frac{77}{97}\times\frac{76}{96}=1-0.319=\mathbf{0.681}$$

$$\beta=P(H_0\text{ を採択}\mid H_1\text{ が真})=P(X=5\mid \text{白球 50 個, 黒球 50 個})$$
$$=\frac{50}{100}\times\frac{49}{99}\times\frac{48}{98}\times\frac{47}{97}\times\frac{46}{96}=\mathbf{0.028}$$

8.13 最初の問題は，第1種の過誤の確率を求める問題である．25個の標本の黒球の数を X とすると，仮説 $H_0: p=\frac{1}{2}$ が真の下で，X は平均 $12.5\left(=25\times\frac{1}{2}\right)$，分散 $6.25\left(=25\times\frac{1}{2}\times\frac{1}{2}\right)$ の2項分布に従う．$n=25$ は大きいから，2項分布への正規近似を使う．

$z=\dfrac{X-12.5}{\sqrt{6.25}}$ とすると，求める確率 α は，

$$\alpha=P(15\leqq X\leqq 20\mid H_0)$$
$$=P\left(\frac{14.5-12.5}{\sqrt{6.25}}<z<\frac{20.5-12.5}{\sqrt{6.25}}\mid H_0\right)$$
$$=P(0.8<z<3.2\mid H_0)=0.211$$

また，後半は，仮説 $H_0:\ p=\frac{1}{2}$, $H_1:\ p>\frac{1}{2}$ の片側検定である．$\alpha=0.05$, $z_{0.05}=1.645$ より，$np=64\times\frac{1}{2}=32$, $np(1-p)=64\times\frac{1}{2}\times\frac{1}{2}=16$ を用いて

$$z=\frac{Y-32}{4}\geqq 1.645$$

よって求める棄却域は $Y\geqq$ **38.58**.

9 章

9.1 帰無仮説は，「各目の出る確率は等しい」ということだから H_0: $p_i=\frac{1}{10}$ ($i=0, 1, 2, \cdots, 9$) で表される．H_0 の下での各期待度数は $200\times\frac{1}{2}=20$ となるから

$$\chi^2=\frac{(26-20)^2}{20}+\frac{(27-20)^2}{20}+\cdots+\frac{(15-20)^2}{20}=11.8$$

χ^2 分布表より $\chi^2_{0.05}(9)=16.92$．棄却域は $\chi^2>16.92$ で，11.8 は棄却域に入らないから，帰無仮説は棄却されない．この乱数サイズは**正しいと判断される**．

9.2 仮説は「0 から 9 までの数字の出現確率は $\frac{1}{10}$ である」ということであるから，この仮説の下では各数に対する期待度数は $100\times\frac{1}{10}=10$ となる．観測度数と期待度数から，χ^2 の値は，6.0（下表より）．

x_i	n_i	m_i	n_i-m_i	$\frac{(n_i-m_i)^2}{m_i}$
0	11	10	1	0.1
1	8	10	−2	0.4
2	13	10	3	0.9
3	12	10	2	0.4
4	6	10	−4	1.6
5	13	10	3	0.9
6	11	10	1	0.1
7	6	10	−4	1.6
8	10	10	0	0
9	10	10	0	0
計	100			$\chi^2=6.0$

$\alpha=0.05$ と $\nu=10-1=9$ より，$\chi^2_{0.05}(9)=16.92$．よって，$\chi^2=6.0$ は有意ではなく，与えられた 100 個の数字は**乱数の一部であるとみなされる**．

9.3 死者の総数 $=1\times 65+2\times 22+3\times 3+4\times 1=122$．よって，(1 兵団 1 年間当たり死者の平均数) $=\frac{122}{200}=0.61$．それゆえ，馬にけられて死んだ兵士の数の分布は $\lambda=0.61$ のポアソン分布 $P(x)=\frac{0.61^x e^{-0.61}}{x!}$ ($x=0, 1, 2, \cdots$) に従うという仮説を立て，その下で期待度数を計算する．

値 x	0	1	2	3	4
確率 $P(x)$	0.543	0.331	0.101	0.021	0.003
期待度数 $200\times P(x)$	109.0	66.3	20.2	4.1	0.6
				4.7	
観測度数	109	65	22	3	1
				4	

$x=3$ と $x=4$ の度数を合併して，χ^2 の値を求めると，

$$\chi^2=\frac{(109-109.0)^2}{109}+\frac{(65-66.3)^2}{66.3}+\frac{(22-20.2)^2}{20.2}+\frac{(4-4.7)^2}{4.7}=0.22$$

λ の推定値を使ったので，自由度は $\nu=4-1-1=2$. $\alpha=0.05$ とすれば，$\chi^2_{0.05}(2)=5.99$. よって，$\chi^2=0.22$ は有意でない．したがって，馬にけられて死んだ兵士の数の分布は**ポアソン分布に従うとみなされる**.

9.4 注射をする，しないということと，流感にかかる，かからないということとは独立であるとの仮説を立てて，独立性の検定を行う．このため，周辺和を用いて，期待度数を計算する．

$$m_{11}=\frac{160\times 76}{260}=46.8,\qquad m_{12}=\frac{100\times 76}{260}=29.2,$$

$$m_{21}=\frac{160\times 184}{260}=113,2,\qquad m_{22}=\frac{100\times 184}{260}=70.8$$

2×2 分割表だから，イエーツの補正を施してから χ^2 の値を求める．

$$\chi^2=\sum\frac{\left(|n_i-m_i|-\dfrac{1}{2}\right)^2}{m_i}=\frac{14.3^2}{46.8}+\frac{14.3^2}{29.2}+\frac{14.3^2}{113.2}+\frac{14.3^2}{70.8}=16.1>3.84=\chi^2_{0.05}(1)$$

よって，仮説は棄却される．したがって，予防注射は**効果があったといえる**.

9.5 周辺和から各枡の期待度数を求めると，

8.67	10.05	7.28
22.33	25.91	18.76
11.00	12.76	9.24
8.00	9.28	6.72

これより

$$\chi^2=\frac{(8-8.67)^2}{8.67}+\cdots+\frac{(12-6.72)^2}{6.72}=16.39$$

$\alpha=0.05$, $\nu=3\times 2=6$ の χ^2 分布表より $\chi^2_{0.05}(6)=12.59$. よって $\chi^2=16.39$ は有意である．したがって学部間に**優劣が認められる**.

10 章

10.1 $n=10$, $\sum x_i=627$, $\sum y_i=631$, $\sum x_i{}^2=43391$, $\sum y_i{}^2=43281$, $\sum x_i y_i=42828$. よって,

$$r=\frac{n\sum x_i y_i-(\sum x_i)(\sum y_i)}{\sqrt{\{n\sum x_i{}^2-(\sum x_i)^2\}\{n\sum y_i{}^2-(\sum y_i)^2\}}}$$

$$=\frac{10\times 42828-627\times 631}{\sqrt{10\times 43391-627^2}\sqrt{10\times 43281-631^2}}=\mathbf{0.86\cancel{8}\,7}$$

10 人の生徒の化学と物理の得点を順位になおし, r_S を求める.

	A	B	C	D	E	F	G	H	I	J
化学の順位	2	1	3	4	5	6	8	7	10	9
物理の順位	3	2	1	4	5.5	7	5.5	9	10	8
d	-1	-1	2	0	-0.5	-1	2.5	-2	0	1
d^2	1	1	4	0	2.5	1	6.25	4	0	1

$$\sum d_i{}^2=20.75$$

$$r_s=1-\frac{6\times 20.75}{10\times(10^2-1)}=\mathbf{0.874}$$

10.2 まず X の平均と分散は

$$\bar{x}=\frac{1}{n}\sum x_i=\frac{1}{n}\{p(a+c)+q(b+d)\}$$

$$s_x{}^2=\frac{1}{n}\sum x_i{}^2-\bar{x}^2=\frac{1}{n}\{p^2(a+c)+q^2(b+d)\}-\bar{x}^2$$

これを整理すると,

$$s_x{}^2=\frac{1}{n^2}(p-q)^2(a+c)(b+d)$$

同様に Y について計算すると,

$$\bar{y}=\frac{1}{n}\{r(a+b)+s(c+d)\}$$

$$s_y{}^2=\frac{1}{n}\{r^2(a+b)+s^2(c+d)\}-\bar{y}^2$$

$$=\frac{1}{n^2}(r-s)^2(a+b)(c+d)$$

さらに共分散は

$$s_{xy}=\frac{1}{n}\sum x_i y_i-\bar{x}\,\bar{y}=\frac{1}{n}\{pra+qrb+psc+qsd\}-\bar{x}\,\bar{y}$$

$$=\frac{1}{n^2}(ad-bc)(p-q)(r-s)$$

したがって相関係数は

$$r=\frac{s_{xy}}{s_x s_y}=\boldsymbol{\frac{ad-bc}{\sqrt{(a+b)(a+c)(b+d)(c+d)}}}$$

10.3 点数データであるが，これを順位データになおす．もし同順位があれば，それら順位の平均値を当てる．

	①	②	③	④	⑤	計
A	5	1	2	3	4	15
B	4	1.5	1.5	3	5	15
$d=A-B$	1	−0.5	0.5	0	−1	0
d^2	1	0.25	0.25	0	1	2.5

$$r=1-\frac{6\times 2.5}{5(5^2-1)}=0.875$$

10.4 変量の値は 1 から n までの自然数だから，

$$\sum_{i=1}^{n}x_i=\sum_{i=1}^{n}y_i=\sum_{i=1}^{n}i=\frac{1}{2}n(n+1)$$

$$\sum_{i=1}^{n}x_i^2=\sum_{i=1}^{n}y_i^2=\sum_{i=1}^{n}i^2=\frac{1}{6}n(n+1)(2n+1)$$

したがって

$$\bar{x}=\frac{1}{2}(n+1)$$

$$s_x^2=\frac{1}{n}\sum x_i^2-\left(\frac{1}{n}\sum x_i\right)^2=\frac{1}{6}(n+1)(2n+1)-\left(\frac{n+1}{2}\right)^2=\frac{1}{12}(n^2-1)$$

同様に $\bar{y}=\frac{1}{2}(n+1)$，$s_y^2=\frac{1}{12}(n^2-1)$ が成り立つ．

$d_i=x_i-y_i$ より両辺を 2 乗して加えれば，

$$\sum d_i^2=\sum x_i^2-2\sum x_iy_i+\sum y_i^2$$

より

$$\sum x_iy_i=\frac{1}{2}\{\sum x_i^2+\sum y_i^2-\sum d_i^2\}$$

であるから

$$\begin{aligned}s_{xy}&=\frac{1}{n}\sum x_iy_i-\bar{x}\bar{y}=\frac{1}{2n}\{\sum x_i^2+\sum y_i^2-\sum d_i^2\}-\bar{x}\bar{y}\\&=\frac{(n+1)(2n+1)}{6}-\frac{(n+1)^2}{4}-\frac{1}{2n}\sum d_i^2\\&=\frac{1}{12}(n^2-1)-\frac{1}{2n}\sum d_i^2\end{aligned}$$

よって

$$r_S=\frac{s_{xy}}{s_xs_y}=\frac{\frac{1}{12}(n^2-1)-\frac{1}{2n}\sum d_i^2}{\frac{1}{12}(n^2-1)}=1-\frac{6\sum_{i=1}^{n}d_i^2}{n(n^2-1)}$$

10.5　審査員1と審査員2が与えた順位の差を d_{12}, 順位相関係数を r_{12} とすれば, 計算より $\sum d_{12}{}^2=84$. よって, $r_{12}=1-\dfrac{6\times 84}{8(8^2-1)}=0$.

同様にして, $\sum d_{23}{}^2=106$, $\sum d_{13}{}^2=58$ より

$$r_{23}=1-\frac{6\times 106}{8(8^2-1)}=-0.26$$

$$r_{13}=-\frac{6\times 58}{8(8^2-1)}=0.31$$

順位の合計が小さい順に決めればよいから, **第1位はF**, **第2位はB**, **第3位はH** と判定される.

10.6　標本相関係数 r を求める. $n=10$, $\sum x_i=557$, $\sum x_i{}^2=32399$, $\sum y_i=656$, $\sum y_i{}^2=44706$, $\sum x_iy_i=37861$ よって,

$$r=\frac{10\times 37861-557\times 656}{\sqrt{10\times 32399-557^2}\sqrt{10\times 44706-656^2}}=0.872$$

小標本による $H_0:\rho=0$, $H_1:\rho\neq 0$ の検定の問題であるから, 統計量

$$t=\sqrt{\frac{r^2(n-2)}{1-r^2}}$$

が H_0 の下で自由度 $n-2$ の t 分布に従うことを使う.

$n=10$, $\alpha=0.01$ に対する棄却域は, $t_{0.005}(8)=3.36$ より $|t|>3.36$.

$$t=\sqrt{\frac{0.872^2(10-2)}{1-0.872^2}}\fallingdotseq 5.04$$

は棄却域に入るから, $H_0:\rho=0$ は棄却される. 実際, 国語と英語の相関係は $r=0.872$ と高く, $\rho=0$ かどうかの検定にかけるまでもない.

10.7　(a)　与えられたデータより, $n=6$, $\sum x_i=90$, $\sum y_i=135.7$, $\sum x_i{}^2=1420$, $\sum x_iy_i=2069.8$ であるから,

$$b=\frac{6\times 2069.8-90\times 135.7}{6\times 1420-90^2}=0.49$$

$$y=\bar{y}+b(x-\bar{x})$$

に代入すると,

$$\boldsymbol{y=15.25+0.49x}$$

(b)　$x=15$ とすると, $y=15.25+0.49\times 15=\mathbf{22.6}$ (cm)

10.8　(a)　表のデータ値より $n=6$ 回測定していて

$$\sum_{i=1}^{6}x_i=2\times(20+25+30)=150,\qquad \sum_{i=1}^{6}y_i=964.4,$$

$$\sum_{i=1}^{6}x_i{}^2=2\times(20^2+25^2+30^2)=3850,\qquad \sum_{i=1}^{6}y_i{}^2=156608.22,$$

$$\sum_{i=1}^{6}x_iy_i=24509$$

これより

$$b=\frac{6\times 24509-150\times 964.4}{6\times 3850-150^2}=3.99$$

$$a=\bar{y}-b\bar{x}=160.73-3.99\times 25=60.98$$

よって求める回帰直線は

$$\boldsymbol{y=60.98+3.99x}$$

（b） 誤差分散 σ^2 の推定値は

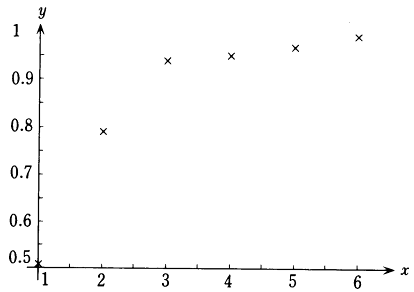
$$s_e=\sqrt{\frac{1}{n-2}\Sigma(y_i-y_i')^2}$$
$$=\sqrt{\frac{1}{n-2}(\Sigma y_i{}^2-a\Sigma y_i-b\Sigma x_iy_i)}$$
$$=\sqrt{\frac{1}{4}(156608.22-60.98\times 964.4-3.99\times 24509)}$$
$$=1.43$$

$n=6$, $\alpha=0.05$, $t_{0.025}(4)=2.78$, $s_e=1.43$. よって β の 95%信頼限界は,

$$b\pm t_{0.025}(5)\frac{s_e}{\sqrt{\Sigma x_i^2-\frac{1}{n}(\Sigma x_i)^2}}=3.99\pm 2.78\times\frac{1.43}{\sqrt{3\,850-\frac{150^2}{6}}}$$
$$=3.99\pm 0.40$$

ゆえに β の信頼区間は(**3.59, 4.39**)

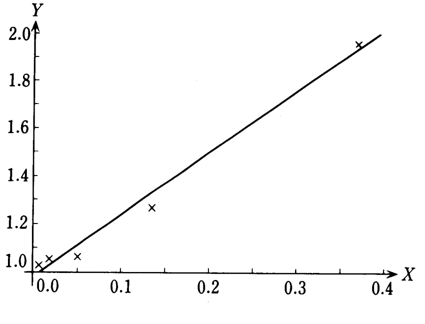
10.9



変換した値を表にすると,

X	0.368	0.135	0.050	0.018	0.007	0.002
Y	1.961	1.266	1.064	1.053	1.031	1.010

これから (X, Y) の散布図は



図より，X と Y の間には直線的傾向がみられるから，最小2乗法で回帰直線を求めると，

$$\boldsymbol{Y=0.98+2.59X}$$

となる．これを近似的に

$$Y=1.0+2.6X$$

とみなせば，もとの x と y の関係は

$$\boldsymbol{y=\frac{1}{1+2.6e^{-x}}}$$

が予想される．この式は，人口増加の研究などによく用いられるロジスティック曲線の1つである．

10.10 与えられたデータから

$$n=5,\ \sum x_i=27,\ \sum y_i=24,\ \sum x_i^2=171,\ \sum y_i^2=134,\ \sum x_iy_i=109$$

Y の X への回帰係数をb, X の Y への回帰係数を b' とすれば

$$b=\frac{n\sum x_iy_i-(\sum x_i)(\sum y_i)}{n\sum x_i^2-(\sum x_i)^2}=\frac{5\times109-27\times24}{5\times171-27^2}=-0.82$$

$$b'=\frac{n\sum x_iy_i-(\sum x_i)(\sum y_i)}{n\sum y_i^2-(\sum y_i)^2}=\frac{5\times109-27\times24}{5\times134-24^2}=-1.10$$

よって，Y の X への回帰直線は，$y-\bar{y}=b(x-\bar{x})$ より

$$y-4.8=-0.82(x-5.4)$$

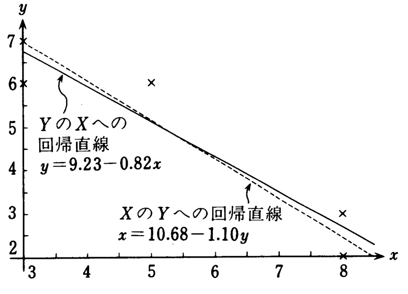
$$\therefore\quad \boldsymbol{y=9.23-0.82x}$$

同様にして，X の Y への回帰直線は，

$$\boldsymbol{x=10.68-1.10y}$$

適当な直線は X の Y への回帰直線であるから，$y=5$ を代入して

$$x=10.68-1.10\times5$$
$$=5.18$$



付　表

付表 **1**　乱 数 表
付表 **2**　正規分布表
付表 **3**　正規分布のパーセント点
付表 **4**　t 分布のパーセント点
付表 **5**　χ^2 分布のパーセント点
付表 **6**　F 分布の 5％ 点
付表 **7**　F 分布の 2.5％ 点

付表 1　乱数表

20832	96698	47328	20139	49458	52617	66677	63268	29060	49358
82315	61278	12763	27543	72365	79797	03693	77801	64956	34446
54976	40751	38209	90857	61735	07022	20252	64647	01093	17240
89212	21445	52181	62029	73208	76879	76119	82012	50368	96570
76034	67815	48817	96142	05048	60063	46681	41221	66292	49382
72085	47435	12885	76397	03150	80407	47957	31730	68007	55760
24628	56011	44777	39122	46028	39860	61586	46117	65279	23905
96561	42369	85507	98773	77822	43504	59835	05081	83499	15146
91391	33464	41713	72924	33296	24540	30592	82458	88681	68938
78262	59374	10663	07279	62844	69298	02374	30192	12464	27854
16938	78301	05249	00668	57479	42228	69326	03366	77120	62603
82815	01572	78988	30045	73843	10910	16205	85186	20533	97009
91903	18644	51015	75488	59200	71050	43410	11972	21223	33024
25843	22092	31098	45195	76443	71471	91952	98187	23328	75732
56905	32621	44630	00497	29084	39128	68471	61400	61617	58330
72976	24058	57625	80853	86270	04099	70235	47319	86477	67150
02570	48357	01721	28831	07758	24589	10132	54772	88925	66643
39833	60594	99191	15140	68947	11924	41682	60710	25599	24387
69527	43974	87918	63610	85849	55559	84024	45213	43770	36088
92044	34825	46419	76189	40101	48070	46919	16483	06323	50572
48396	47600	18836	57958	03974	10160	55763	35850	16779	94794
45228	99878	39937	42119	65358	15659	76571	42766	44276	98828
79806	37358	60109	15000	52767	16520	40980	79809	48579	20878
65017	58873	70370	41057	60341	67821	71261	17679	05078	72277
54378	41374	27357	87866	72848	52763	05298	80212	29792	42473
67031	64940	78341	51128	90677	07675	21613	69353	04358	24046
12738	37773	86209	79675	54842	47014	64345	90185	01044	37700
16895	21201	54476	00455	71987	88354	68258	75909	07739	57262
45515	54680	52284	43553	39375	77398	83743	12851	52962	34684
30238	80786	39399	67168	69900	12973	01813	65471	30848	25047
93332	70220	91214	82629	17072	72039	79116	01340	26167	11551
72683	46812	23737	79388	00037	34666	62911	16166	39308	30528
46809	12516	18616	50023	28559	09061	16089	58776	11286	96429
59853	72411	48115	15238	28028	56554	42170	56120	48744	26829
46577	59697	00122	48863	64461	16593	10291	38279	48946	67437

付表 2　正規分布表

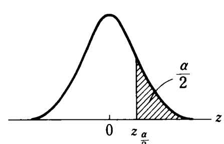
$z \to \Phi(z)$

z	$\Phi(z)$		z	$\Phi(z)$		z	$\Phi(z)$		z	$\Phi(z)$		z	$\Phi(z)$	
0.00	0.5000	40	0.50	0.6915	35	1.00	0.8413	25	1.50	0.9332	13	2.00	0.97725	53
.01	.5040	40	.51	.6950	35	.01	.8438	23	.51	.9345	12	.01	.97778	53
.02	.5080	40	.52	.6985	34	.02	.8461	24	.52	.9357	13	.02	.97831	51
.03	.5120	40	.53	.7019	35	.03	.8485	23	.53	.9370	12	.03	.97882	50
.04	.5160	39	.54	.7054	34	.04	.8508	23	.54	.9382	12	.04	.97932	50
.05	.5199	40	.55	.7088	35	.05	.8531	23	.55	.9394	12	.05	.97982	48
.06	.5239	40	.56	.7123	34	.06	.8554	23	.56	.9406	12	.06	.98030	47
.07	.5279	40	.57	.7157	33	.07	.8577	22	.57	.9418	11	.07	.98077	47
.08	.5319	40	.58	.7190	34	.08	.8599	22	.58	.9429	12	.08	.98124	45
.09	.5359	39	.59	.7224	33	.09	.8621	22	.59	.9441	11	.09	.98169	45
0.10	0.5398	40	0.60	0.7257	34	1.10	0.8643	22	1.60	0.9452	11	2.10	0.98214	43
.11	.5438	40	.61	.7291	33	.11	.8665	21	.61	.9463	11	.11	.98257	43
.12	.5478	39	.62	.7324	33	.12	.8686	22	.62	.9474	10	.12	.98300	41
.13	.5517	40	.63	.7357	32	.13	.8708	21	.63	.9484	11	.13	.98341	41
.14	.5557	39	.64	.7389	33	.14	.8729	20	.64	.9495	10	.14	.98382	40
.15	.5596	40	.65	.7422	32	.15	.8749	21	.65	.9505	10	.15	.98422	39
.16	.5636	39	.66	.7454	32	.16	.8770	20	.66	.9515	10	.16	.98461	39
.17	.5675	39	.67	.7486	31	.17	.8790	20	.67	.9525	10	.17	.98500	37
.18	.5714	39	.68	.7517	32	.18	.8810	20	.68	.9535	10	.18	.98537	37
.19	.5753	40	.69	.7549	31	.19	.8830	19	.69	.9545	9	.19	.98574	36
0.20	0.5793	39	0.70	0.7580	31	1.20	0.8849	20	1.70	0.9554	10	2.20	0.98610	35
.21	.5832	39	.71	.7611	31	.21	.8869	19	.71	.9564	9	.21	.98645	34
.22	.5871	39	.72	.7642	31	.22	.8888	19	.72	.9573	9	.22	.98679	34
.23	.5910	38	.73	.7673	31	.23	.8907	18	.73	.9582	9	.23	.98713	32
.24	.5948	39	.74	.7704	30	.24	.8925	19	.74	.9591	8	.24	.98745	33
.25	.5987	39	.75	.7734	30	.25	.8944	18	.75	.9599	9	.25	.98778	31
.26	.6026	38	.76	.7764	30	.26	.8962	18	.76	.9608	8	.26	.98809	31
.27	.6064	39	.77	.7794	29	.27	.8980	17	.77	.9616	9	.27	.98840	30
.28	.6103	38	.78	.7823	29	.28	.8997	18	.78	.9625	8	.28	.98870	29
.29	.6141	38	.79	.7852	29	.29	.9015	17	.79	.9633	8	.29	.98899	29
0.30	0.6179	38	0.80	0.7881	29	1.30	0.9032	17	1.80	0.9641	8	2.30	0.98928	28
.31	.6217	38	.81	.7910	29	.31	.9049	17	.81	.9649	7	.31	.98956	27
.32	.6255	38	.82	.7939	28	.32	.9066	16	.82	.9656	8	.32	.98983	27
.33	.6293	38	.83	.7967	28	.33	.9082	17	.83	.9664	7	.33	.99010	26
.34	.6331	37	.84	.7995	28	.34	.9099	16	.84	.9671	7	.34	.99036	25
.35	.6368	38	.85	.8023	28	.35	.9115	16	.85	.9678	8	.35	.99061	25
.36	.6406	37	.86	.8051	27	.36	.9131	16	.86	.9686	7	.36	.99086	25
.37	.6443	37	.87	.8078	28	.37	.9147	15	.87	.9693	6	.37	.99111	23
.38	.6480	37	.88	.8106	27	.38	.9162	15	.88	.9699	7	.38	.99134	24
.39	.6517	37	.89	.8133	26	.39	.9177	15	.89	.9706	7	.39	.99158	22
0.40	.6554	37	0.90	0.8159	27	1.40	0.9192	15	1.90	0.9713	6	2.40	0.99180	22
.41	.6591	37	.91	.8186	26	.41	.9207	15	.91	.9719	7	.41	.99202	22
.42	.6628	36	.92	.8212	26	.42	.9222	14	.92	.9726	6	.42	.99224	21
.43	.6664	36	.93	.8238	26	.43	.9236	15	.93	.9732	6	.43	.99245	21
.44	.6700	36	.94	.8264	25	.44	.9251	14	.94	.9738	6	.44	.99266	20
.45	.6736	36	.95	.8289	26	.45	.9265	14	.95	.9744	6	.45	.99286	19
.46	.6772	36	.96	.8315	25	.46	.9279	13	.96	.9750	6	.46	.99305	19
.47	.6808	36	.97	.8340	25	.47	.9292	14	.97	.9756	5	.47	.99324	19
.48	.6844	35	.98	.8365	24	.48	.9306	13	.98	.9761	6	.48	.99343	18
.49	.6879	36	.99	.8389	24	.49	.9319	13	.99	.9767	5	.49	.99361	18

正規分布表(続き)

z	$\Phi(z)$		z	$\Phi(z)$		z	$\Phi(z)$	
2.50	0.99379	17	2.70	0.99653	11	2.90	0.99813	6
.51	.99396	17	.71	.99664	10	.91	.99819	6
.52	.99413	17	.72	.99674	9	.92	.99825	6
.53	.99430	16	.73	.99683	10	.93	.99831	5
.54	.99446	15	.74	.99693	9	.94	.99836	5
.55	.99461	16	.75	.99702	9	.95	.99841	5
.56	.99477	15	.76	.99711	9	.96	.99846	5
.57	.99492	14	.77	.99720	8	.97	.99851	5
.58	.99506	14	.78	.99728	8	.98	.99856	5
.59	.99520	14	.79	.99736	8	.99	.99861	4
2.60	0.99534	13	2.80	0.99744	8	3.0	0.99865	38
.61	.99547	13	.81	.99752	8	3.1	.99903	28
.62	.99560	13	.82	.99760	7	3.2	.99931	21
.63	.99573	12	.83	.99767	7	3.3	.99952	14
.64	.99585	13	.84	.99774	7	3.4	.99966	11
.65	.99598	11	.85	.99781	7	3.5	.99977	7
.66	.99609	12	.86	.99788	7	3.6	.99984	5
.67	.99621	11	.87	.99795	6	3.7	.99989	4
.68	.99632	11	.88	.99801	6	3.8	.99993	2
.69	.99643	10	.89	.99807	6	3.9	.99995	2
						4.0	0.99997	

付表 3　正規分布のパーセント点



$\alpha \rightarrow z_{\frac{\alpha}{2}}$

α	0.20	0.10	0.05	0.02	0.01	0.002
$\frac{\alpha}{2}$	0.10	0.05	0.025	0.01	0.005	0.001
$z_{\frac{\alpha}{2}}$	1.282	1.645	1.960	2.326	2.576	3.090

付表 4　t 分布のパーセント点

$\alpha, \nu \to t_{\frac{\alpha}{2}}(\nu)$

ν \ α	0.20	0.10	0.05	0.02	0.01
$\frac{\alpha}{2}$	0.10	0.05	0.025	0.01	0.005
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.179	2.681	3.055
13	1.350	1.771	2.160	2.650	3.012
14	1.345	1.761	2.145	2.624	2.977
15	1.341	1.753	2.131	2.602	2.947
16	1.337	1.746	2.120	2.583	2.921
17	1.333	1.740	2.110	2.567	2.898
18	1.330	1.734	2.101	2.552	2.878
19	1.328	1.729	2.093	2.539	2.861
20	1.325	1.725	2.086	2.528	2.845
21	1.323	1.721	2.080	2.518	2.831
22	1.321	1.717	2.074	2.508	2.819
23	1.319	1.714	2.069	2.500	2.807
24	1.318	1.711	2.064	2.492	2.797
25	1.316	1.708	2.060	2.485	2.787
26	1.315	1.706	2.056	2.479	2.779
27	1.314	1.703	2.052	2.473	2.771
28	1.313	1.701	2.048	2.467	2.763
29	1.311	1.699	2.045	2.462	2.756
30	1.310	1.697	2.042	2.457	2.750
40	1.303	1.684	2.021	2.423	2.704
60	1.296	1.671	2.000	2.390	2.660
120	1.289	1.658	1.980	2.358	2.617
∞	1.282	1.645	1.960	2.326	2.576

付表 5　χ^2 分布のパーセント点

$\alpha, \nu \to \chi^2_\alpha(\nu)$

ν \ α	0.995	0.975	0.05	0.025	0.01	0.005
1	0.0000393	0.000982	3.84	5.02	6.63	7.88
2	0.0100	0.0506	5.99	7.38	9.21	10.60
3	0.0717	0.216	7.81	9.35	11.34	12.84
4	0.207	0.484	9.49	11.14	13.28	14.86
5	0.412	0.831	11.07	12.83	15.09	16.75
6	0.676	1.24	12.59	14.45	16.81	18.55
7	0.989	1.69	14.07	16.01	18.48	20.28
8	1.34	2.18	15.51	17.53	20.09	21.95
9	1.73	2.70	16.92	19.02	21.67	23.59
10	2.16	3.25	18.31	20.48	23.21	25.19
11	2.60	3.82	19.68	21.92	24.73	26.76
12	3.07	4.40	21.03	23.34	26.22	28.30
13	3.57	5.01	22.36	24.74	27.69	29.82
14	4.07	5.63	23.68	26.12	29.14	31.32
15	4.60	6.26	25.00	27.49	30.58	32.80
16	5.14	6.91	26.30	28.85	32.00	34.27
17	5.70	7.56	27.59	30.19	33.41	35.72
18	6.26	8.23	28.87	31.53	34.81	37.16
19	6.84	8.91	30.14	32.85	36.19	38.58
20	7.43	9.59	31.41	34.17	37.57	40.00
21	8.03	10.28	32.67	35.48	38.93	41.40
22	8.64	10.98	33.92	36.78	40.29	42.80
23	9.26	11.69	35.17	38.08	41.64	44.18
24	9.89	12.40	36.42	39.36	42.98	45.56
25	10.52	13.12	37.65	40.65	44.31	46.93
26	11.16	13.84	38.89	41.92	45.64	48.29
27	11.81	14.57	40.11	43.19	46.96	49.64
28	12.46	15.31	41.34	44.46	48.28	50.99
29	13.12	16.05	42.56	45.72	49.59	52.34
30	13.79	16.79	43.77	46.98	50.89	53.67
40	20.71	24.43	55.76	59.34	63.69	66.77
50	27.99	32.36	67.50	71.42	76.15	79.49
60	35.53	40.48	79.08	83.30	88.38	91.95
70	43.28	48.76	90.53	95.02	100.4	104.2
80	51.17	57.15	101.9	106.6	112.3	116.3
90	59.20	65.65	113.1	118.1	124.1	128.3
100	67.33	74.22	124.3	129.6	135.8	140.2

付表 6　F 分布の 5% 点

$\nu_1,\ \nu_2 \to F_{0.05}(\nu_1,\ \nu_2)$

ν_2 \ ν_1	1	2	3	4	5	6	7	8	10	12	24	∞
1	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	241.9	243.9	249.0	254.3
2	18.5	19.0	19.2	19.2	19.3	19.3	19.4	19.4	19.4	19.4	19.5	19.5
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.79	8.74	8.64	8.53
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	5.96	5.91	5.77	5.63
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.74	4.68	4.53	4.36
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.06	4.00	3.84	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.64	3.57	3.41	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.35	3.28	3.12	2.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.14	3.07	2.90	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	2.98	2.91	2.74	2.54
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.85	2.79	2.61	2.40
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.75	2.69	2.51	2.30
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.67	2.60	2.42	2.21
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.60	2.53	2.35	2.13
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.54	2.48	2.29	2.07
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.49	2.42	2.24	2.01
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.45	2.38	2.19	1.96
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.41	2.34	2.15	1.92
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.38	2.31	2.11	1.88
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.35	2.28	2.08	1.84
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.32	2.25	2.05	1.81
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.30	2.23	2.03	1.78
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.27	2.20	2.00	1.76
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.25	2.18	1.98	1.73
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.24	2.16	1.96	1.71
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.22	2.15	1.95	1.69
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.20	2.13	1.93	1.67
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.19	2.12	1.91	1.65
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.18	2.10	1.90	1.64
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.16	2.09	1.89	1.62
32	4.15	3.29	2.90	2.67	2.51	2.40	2.31	2.24	2.14	2.07	1.86	1.59
34	4.13	3.28	2.88	2.65	2.49	2.38	2.29	2.23	2.12	2.05	1.84	1.57
36	4.11	3.26	2.87	2.63	2.48	2.36	2.28	2.21	2.11	2.03	1.82	1.55
38	4.10	3.24	2.85	2.62	2.46	2.35	2.26	2.19	2.09	2.02	1.81	1.53
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.08	2.00	1.79	1.51
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	1.99	1.92	1.70	1.39
120	3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.91	1.83	1.61	1.25
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.83	1.75	1.52	1.00

F が1より大きくなるようにとり，そのとき，ν_1 が分子の自由度で，ν_2 は分母の自由度である．

付表 7　F 分布の 2.5 % 点

$\nu_1, \nu_2 \rightarrow F_{0.025}(\nu_1, \nu_2)$

ν_2 \ ν_1	1	2	3	4	5	6	7	8	10	12	24	∞
1	648	800	864	900	922	937	948	957	969	977	997	1018
2	38.5	39.0	39.2	39.2	39.3	39.3	39.4	39.4	39.4	39.4	39.5	39.5
3	17.4	16.0	15.4	15.1	14.9	14.7	14.6	14.5	14.4	14.3	14.1	13.9
4	12.22	10.65	9.98	9.60	9.36	9.20	9.07	8.98	8.84	8.75	8.51	8.26
5	10.01	8.43	7.76	7.39	7.15	6.98	6.85	6.76	6.62	6.52	6.28	6.02
6	8.81	7.26	6.60	6.23	5.99	5.82	5.70	5.60	5.46	5.37	5.12	4.85
7	8.07	6.54	5.89	5.52	5.29	5.12	4.99	4.90	4.76	4.67	4.42	4.14
8	7.57	6.06	5.42	5.05	4.82	4.65	4.53	4.43	4.30	4.20	3.95	3.67
9	7.21	5.71	5.08	4.72	4.48	4.32	4.20	4.10	3.96	3.87	3.61	3.33
10	6.94	5.46	4.83	4.47	4.24	4.07	3.95	3.85	3.72	3.62	3.37	3.08
11	6.72	5.26	4.63	4.28	4.04	3.88	3.76	3.66	3.53	3.43	3.17	2.88
12	6.55	5.10	4.47	4.12	3.89	3.73	3.61	3.51	3.37	3.28	3.02	2.72
13	6.41	4.97	4.35	4.00	3.77	3.60	3.48	3.39	3.25	3.15	2.89	2.60
14	6.30	4.86	4.24	3.89	3.66	3.50	3.38	3.29	3.15	3.05	2.79	2.49
15	6.20	4.76	4.15	3.80	3.58	3.41	3.29	3.20	3.06	2.96	2.70	2.40
16	6.12	4.69	4.08	3.73	3.50	3.34	3.22	3.12	2.99	2.89	2.63	2.32
17	6.04	4.62	4.01	3.66	3.44	3.28	3.16	3.06	2.92	2.82	2.56	2.25
18	5.98	4.56	3.95	3.61	3.38	3.22	3.10	3.01	2.87	2.77	2.50	2.19
19	5.92	4.51	3.90	3.56	3.33	3.17	3.05	2.96	2.82	2.72	2.45	2.13
20	5.87	4.46	3.86	3.51	3.29	3.13	3.01	2.91	2.77	2.68	2.41	2.09
21	5.83	4.42	3.82	3.48	3.25	3.09	2.97	2.87	2.73	2.64	2.37	2.04
22	5.79	4.38	3.78	3.44	3.22	3.05	2.93	2.84	2.70	2.60	2.33	2.00
23	5.75	4.35	3.75	3.41	3.18	3.02	2.90	2.81	2.67	2.57	2.30	1.97
24	5.72	4.32	3.72	3.38	3.15	2.99	2.87	2.78	2.64	2.54	2.27	1.94
25	5.69	4.29	3.69	3.35	3.13	2.97	2.85	2.75	2.61	2.51	2.24	1.91
26	5.66	4.27	3.67	3.33	3.10	2.94	2.82	2.73	2.59	2.49	2.22	1.88
27	5.63	4.24	3.65	3.31	3.08	2.92	2.80	2.71	2.57	2.47	2.19	1.85
28	5.61	4.22	3.63	3.29	3.06	2.90	2.78	2.69	2.55	2.45	2.17	1.83
29	5.59	4.20	3.61	3.27	3.04	2.88	2.76	2.67	2.53	2.43	2.15	1.81
30	5.57	4.18	3.59	3.25	3.03	2.87	2.75	2.65	2.51	2.41	2.14	1.79
32	5.53	4.15	3.56	3.22	3.00	2.84	2.72	2.62	2.48	2.38	2.10	1.75
34	5.50	4.12	3.53	3.19	2.97	2.81	2.69	2.59	2.45	2.35	2.08	1.72
36	5.47	4.09	3.53	3.17	2.94	2.79	2.66	2.57	2.43	2.33	2.05	1.69
38	5.45	4.07	3.48	3.15	2.92	2.76	2.64	2.55	2.41	2.31	2.03	1.66
40	5.42	4.05	3.46	3.13	2.90	2.74	2.62	2.53	2.39	2.29	2.01	1.64
60	5.29	3.93	3.34	2.01	2.79	2.63	2.51	2.41	2.27	2.17	1.88	1.48
120	5.15	3.80	3.23	2.89	2.67	2.52	2.39	2.30	2.16	2.05	1.76	1.31
∞	5.02	3.69	3.12	2.79	2.57	2.41	2.29	2.19	2.05	1.94	1.64	1.00

F が1より大きくなるようにとり，そのとき，ν_1 が分子の自由度で，ν_2 は分母の自由度である．

索　引

さ 行

た 行

な　行

は　行

著者略歴

村上正康（むらかみ まさやす）

1947年　九州大学理学部数学科卒業
1968年　千葉大学教養部教授
1989年　千葉大学名誉教授

主要著書

入門数理統計学（共訳, 培風館, 1980）
やさしい例による統計入門（上, 下）（監訳, 培風館, 1979, 1980）
初等統計学　原書第4版（共訳, 培風館, 1981）
教養の線形代数　改訂版（共著, 培風館, 1985）
統計学用語辞典（共訳, 丸善, 1987）
統計学の基礎（訳, 培風館, 1988）
演習線形代数　改訂版（共著, 培風館, 1989）

安田正實（やすだ まさみ）

1971年　九州大学大学院理学研究科修士課程修了
現　在　千葉大学理学部教授
　　　　理学博士

主要著書

確率論入門（訳, 東京図書, 1973）
やさしい例による統計入門（上, 下）（共訳, 培風館, 1979, 1980）
教養の微分積分　改訂版（共著, 培風館, 1986）
統計学用語辞典（共訳, 丸善, 1987）

1989年1月15日　初版発行
1995年9月20日　初版第9刷発行

統計学演習

著　者　村上正康
　　　　安田正實
発行者　山本　格

発行所　株式会社　培風館
東京都千代田区九段南4-3-12・郵便番号102
電話(03)3262-5256(代表)・振替00140-7-44725

前田印刷・坂本製本

PRINTED IN JAPAN

ISBN4-563-00870-2 C3033