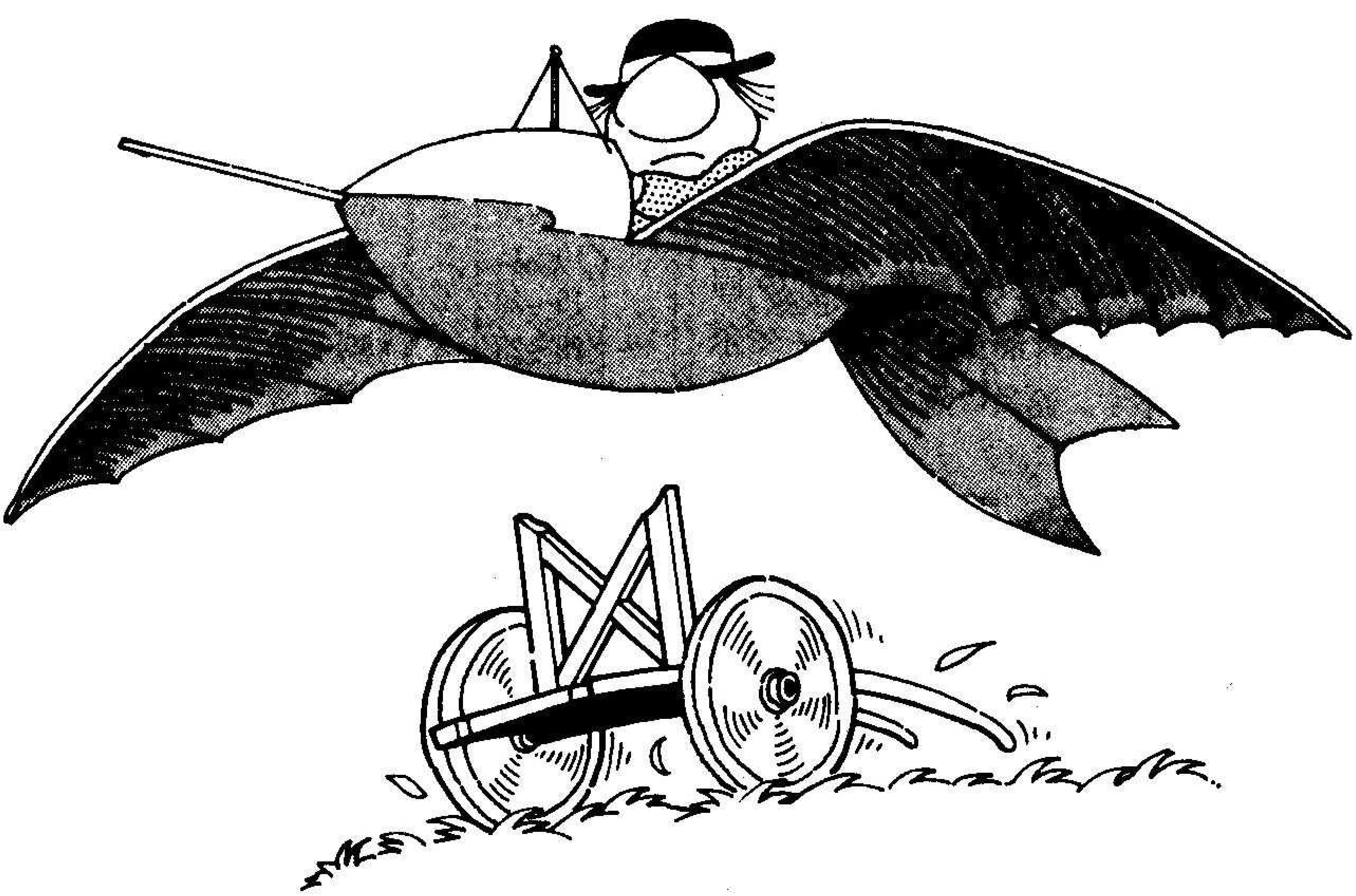
第3章

度数分布と確率分布



○度数分布とは何か

1回目　年　月　日
2回目　年　月　日
3回目　年　月　日

◆われわれが何かをみて，考えるとき，まず，大まかな定義を与える。次に分類する，そのとき，度数分布があらわれる。

◆ さっそくながら，次のデータからはじめるとしましょう。

練習 1. 0から9までの数字をデタラメに書いたところ右のようになった。（こうしてみると，デタラメに書くということが意外とめんどうであることがわかります。もっとも多いのが4の10個，もっとも少ないのが7の2個です）

この度数分布表から相対度数分布表をつくれ。

数字	度数
0	3
1	9
2	9
3	7
4	10
5	7
6	3
7	2
8	4
9	6
計	60

ヒント　度数を全体の個数60で割ったものを相対度数というのですから，下のようになります。

数字	0	1	2	3	4	5	6	7	8	9
相対度数	0.05	0.15	0.15	0.12	0.17	0.12	0.05	0.03	0.07	0.10

注　数をデタラメに配列したものを**乱数表**といいます。p. 194参照。なお，乱数表は心理検査などに利用されています。

練習 2. 下の表は物質の融解熱（1g の固体を液化するに必要な熱量をカロリーで示したもの）を示す。適当に階級に分け，度数分布表をつくれ。

亜鉛	27.0	カリウム	14.4
アンモニア	84	金	15
アルミニウム	95.3	銀	26.5
アンチモン	38.5	氷	79.7
一酸化炭素	7.18	酢酸	47
エチルアルコール	26.1	酸素	3.3
塩化水素	13.05	臭素	16
塩化ナトリウム	116	水銀	2.7
塩素	21.6	水素	13.9
スズ	14.1	ニッケル	71.7
窒素	6.1	白金	24.1
鉄	65	ベンゼン	30.1
銅	49.9	マグネシウム	88
ナトリウム	27	メチルアルコール	23.7
ナフタレン	33.7	リン（黄）	5.1
鉛	5.5	メタン	14.0
二酸化炭素	43.2		

ヒント　これをみると，ずいぶんマチマチです。最小は水銀の2.7，最大は塩化ナトリウムの116です。

資料の数は全部で33個です。

さて，数値の小さい方から並べてみると，右下のようになります。

そこで0～5（0以上5未満），5～10，10～15，……というように階級に分けて度数分布表をつくると左下のようになります。

2.7	26.5
3.3	27.0
5.1	27
5.5	30.1
6.1	33.7
7.18	38.5
13.05	43.2
13.9	47
14.0	49.9
14.1	65
14.4	71.7
15	79.7
16	84
21.6	88
23.7	95.3
24.1	116
26.1	

融解熱	度数	融解熱	度数
0～5	2	60～65	0
5～10	4	65～70	1
10～15	5	70～75	1
15～20	2	75～80	1
20～25	3	80～85	1
25～30	4	85～90	1
30～35	2	90～95	0
35～40	1	95～100	1
40～45	1	100～105	0
45～50	2	105～110	0
50～55	0	110～115	0
55～60	0	115～120	1

さらに，0～10，10～20，……と級間を大きくしてつくってみますと，次のようです。

このように級間を大きくしていくと，だんだん全体の傾向をつかむことができるように

融解熱	度数	融解熱	度数
0～10	6	60～ 70	1
10～20	7	70～ 80	2
20～30	7	80～ 90	2
30～40	3	90～100	1
40～50	3	100～110	0
50～60	0	110～120	1

なってきます。さらに，級間を大きくして0～20，20～40……としますと右のようです。

融解熱	度数
0～ 20	13
20～ 40	10
40～ 60	3
60～ 80	3
80～100	3
100～120	1
計	33

こうしてみると，ここにきて，はじめて，もっとも特徴のよくつかめる度数分布表が得られたのです。もちろん，度数分布表をつくる目的にもよるのですが，数字の分布だけからみるとこの分布表がもっともよいといっていいでしょう。

*　*　*

◆ 同じことながら，次の練習をやってみませんか。

練習 3. 次の資料はある高校1年生女子の胸囲（cm）である。適当に階級に分けて，度数分布表をつくれ。

77.5　92.0　87.5　83.0　79.0　85.0
89.0　78.0　79.0　92.0　80.0　79.0
82.5　80.0　85.5　79.0　79.0　79.0
78.0　78.0　84.5　83.0　83.5　80.0
82.5　83.5　86.5　82.0　80.0　86.5
87.0　79.5　80.0　80.5　80.0　79.0
92.0　82.0

ヒント　これは練習2.とちがって0.5キザミの数値ですから，まず各変量についての度数を調べてみましょう。この際，たてに変量をとり，各値が出てきたら |，||，|||，||||，卌，卌|といったぐあいに記録してゆくのがふつうです。そして，10になったら卌卌としておくのです。こうしてみると，どうやら級間3～5 cmといったところがよさそうです。

<table>
<tr><th>変量</th><th></th><th>度数1</th><th>度数2</th><th>度数3</th><th>度数4</th><th>度数5</th></tr>
<tr><td>77.5</td><td>|</td><td>1</td><td>1</td><td rowspan="5">12</td><td rowspan="5">12</td><td rowspan="5">12</td></tr>
<tr><td>78.0</td><td>|||</td><td>3</td><td rowspan="2">3</td></tr>
<tr><td>78.5</td><td></td><td>0</td></tr>
<tr><td>79.0</td><td>卌||</td><td>7</td><td rowspan="2">8</td></tr>
<tr><td>79.5</td><td>|</td><td>1</td></tr>
<tr><td>80.0</td><td>卌|</td><td>6</td><td rowspan="2">7</td><td rowspan="6">11</td><td rowspan="10">16</td><td rowspan="20">23</td></tr>
<tr><td>80.5</td><td>|</td><td>1</td></tr>
<tr><td>81.0</td><td></td><td>0</td><td rowspan="2">0</td></tr>
<tr><td>81.5</td><td></td><td>0</td></tr>
<tr><td>82.0</td><td>||</td><td>2</td><td rowspan="2">4</td></tr>
<tr><td>82.5</td><td>||</td><td>2</td></tr>
<tr><td>83.0</td><td>||</td><td>2</td><td rowspan="2">4</td><td rowspan="6">7</td></tr>
<tr><td>83.5</td><td>||</td><td>2</td></tr>
<tr><td>84.0</td><td></td><td>0</td><td rowspan="2">1</td></tr>
<tr><td>84.5</td><td>|</td><td>1</td></tr>
<tr><td>85.0</td><td>|</td><td>1</td><td rowspan="2">2</td><td rowspan="10">7</td></tr>
<tr><td>85.5</td><td>|</td><td>1</td></tr>
<tr><td>86.0</td><td></td><td>0</td><td rowspan="2">2</td><td rowspan="6">4</td></tr>
<tr><td>86.5</td><td>||</td><td>2</td></tr>
<tr><td>87.0</td><td>|</td><td>1</td><td rowspan="2">2</td></tr>
<tr><td>87.5</td><td>|</td><td>1</td></tr>
<tr><td>88.0</td><td></td><td>0</td><td rowspan="2">0</td></tr>
<tr><td>88.5</td><td></td><td>0</td></tr>
<tr><td>89.0</td><td>|</td><td>1</td><td rowspan="2">1</td><td rowspan="6">1</td></tr>
<tr><td>89.5</td><td></td><td>0</td></tr>
<tr><td>90.0</td><td></td><td>0</td><td rowspan="2">0</td><td rowspan="6">3</td><td rowspan="6">3</td></tr>
<tr><td>90.5</td><td></td><td>0</td></tr>
<tr><td>91.0</td><td></td><td>0</td><td rowspan="2">0</td></tr>
<tr><td>91.5</td><td></td><td>0</td></tr>
<tr><td>92.0</td><td>|||</td><td>3</td><td rowspan="2">3</td><td rowspan="2">3</td></tr>
<tr><td>92.5</td><td></td><td>0</td></tr>
</table>

1回目　年　月　日
2回目　年　月　日
3回目　年　月　日

○ヒストグラムとは何か

◆数値は正確ですが，直観的ではありません。
これを見易すくするのが図的表現というもの。
ここにヒストグラムが登場する。

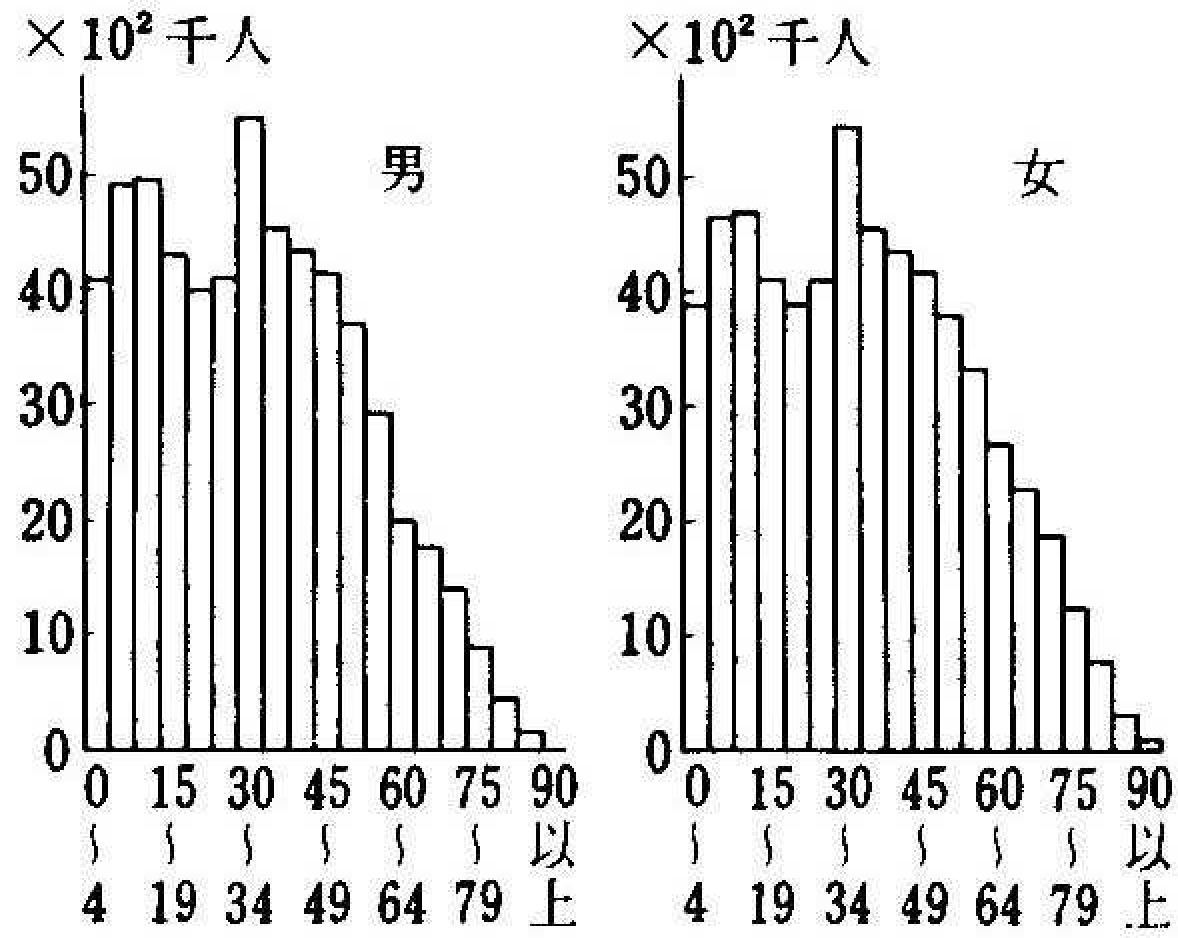
◆ 度数分布を柱状のグラフで表したものを**ヒストグラム**といい，柱の上の辺の中点を順に結んでできるグラフを**度数分布多角形**といいます。

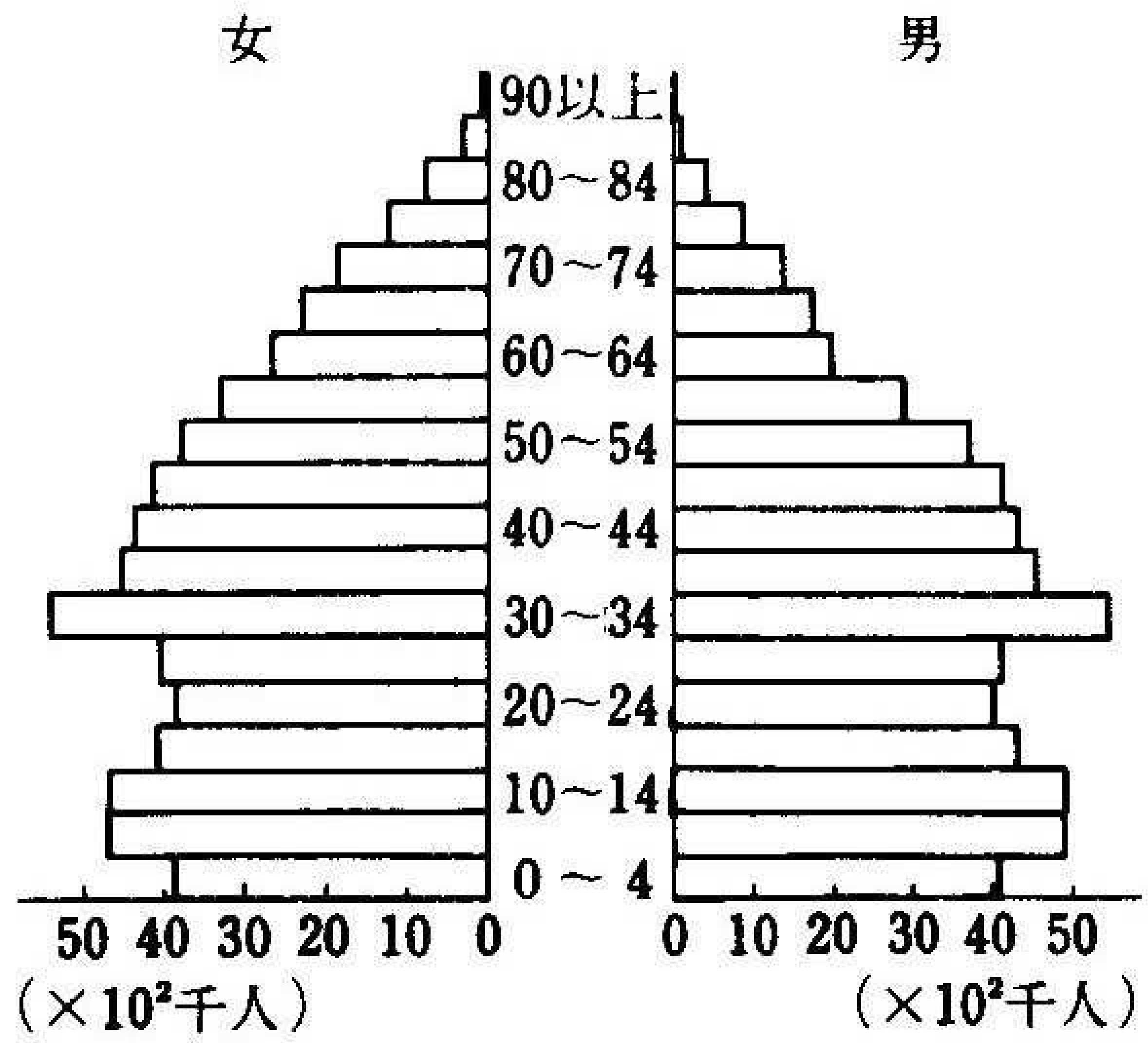
練習 1. 下の表はわが国の年齢，男女別人口である。男女別にヒストグラムをつくれ。

ヒント　もちろん，この数値をそのまま正確にヒストグラムに表すことはできません。30～34歳の人数を基準にとって，これをどのくらいの大きさに表すかが問題でしょう。

なお，人口の場合には，年齢の軸をたてにとって，男女のヒストグラムをその左側と右側に作ったものがよく使われています。それを右に示してあります。

年齢	男	女
	(千人)	(千人)
0～ 4	4090	3881
5～ 9	4928	4686
10～14	4962	4708
15～19	4306	4107
20～24	3997	3885
25～29	4132	4082
30～34	5498	5436
35～39	4563	4547
40～44	4352	4372
45～49	4136	4167
50～54	3740	3805
55～59	2950	3325
60～64	2002	2687
65～69	1769	2289
70～74	1410	1878
75～79	891	1252
80～84	474	769
85～89	162	316
90以上	39	100





＊　＊　＊

上のヒストグラムと下のヒストグラムをみて，何か気のつくことはありませんか。

30～34で男女とも突出しているが，ハテどうしたわけであろうか，とか，55くらいから急に男の方が少なくなるのはどうしてであろうか，など，いろいろ見えてくる!!

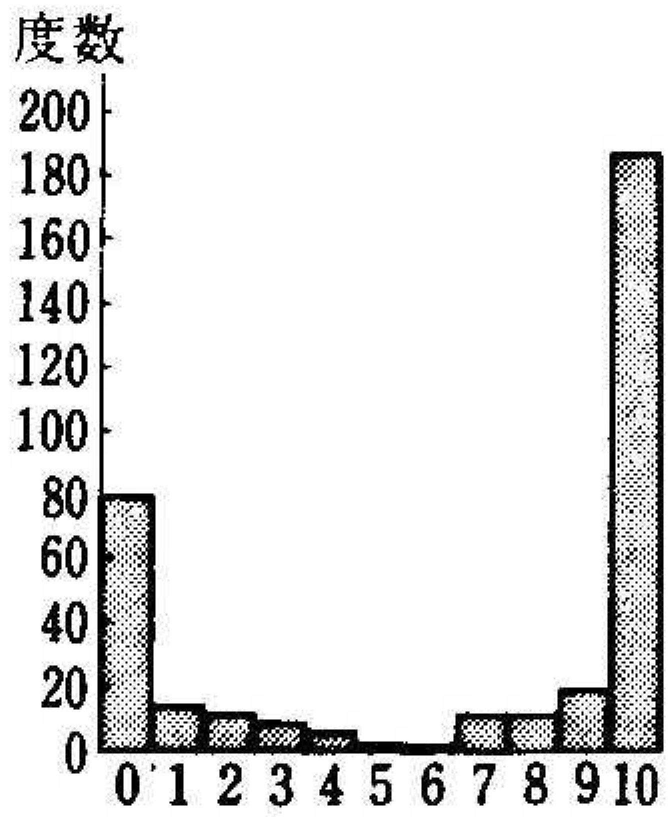
＊　＊　＊

◆ ヒストグラムを使って統計量の意味を考えておきましょう。たとえば，雲の多い少ないを表すのに雲量（ウンリョウ）というものを考えます。これは，全天の雲の占める面積を0から10までで表したものです。

さて，20年間，東京である時刻の雲量を調べたものがありますが，そのヒストグラムを示すと次のような図になります。これによると，もっとも多いのが10で，次が0でもっとも少ないのが5です。つまり，分布がU字形になっています。これを**U字型分布**といいま

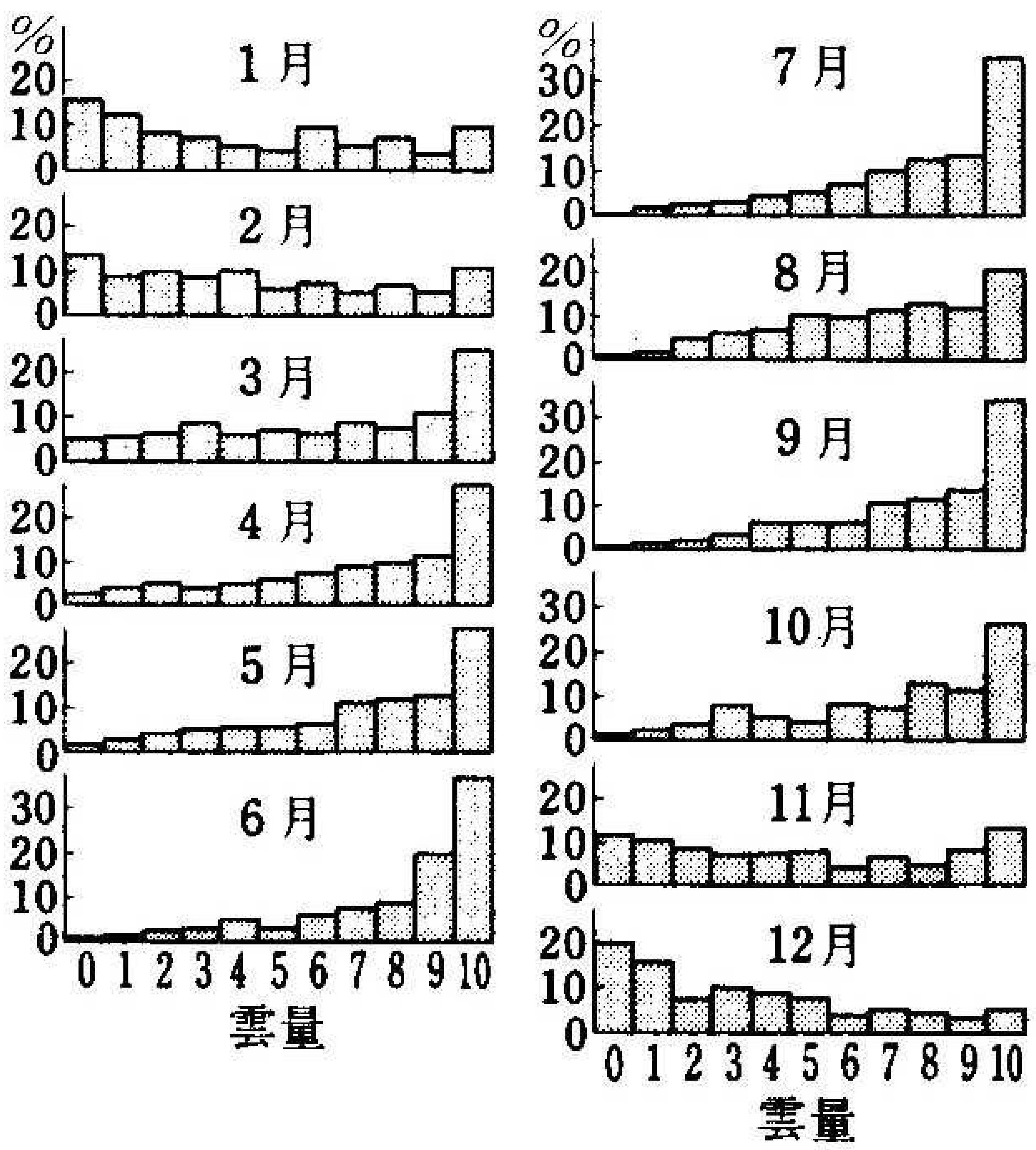
す。ところが，このような雲量も1日平均とか半旬平均となると，まったくちがってくるのです。

下には東京の同じ期間の日平均雲量，半旬平均雲量，旬平均雲量および年平均雲量のヒストグラムを月別に示してありますが，そのちがいがよくあらわれています。（なお，年平均雲量は上の資料とちがって，もっと長い期間のものからつくってあります）

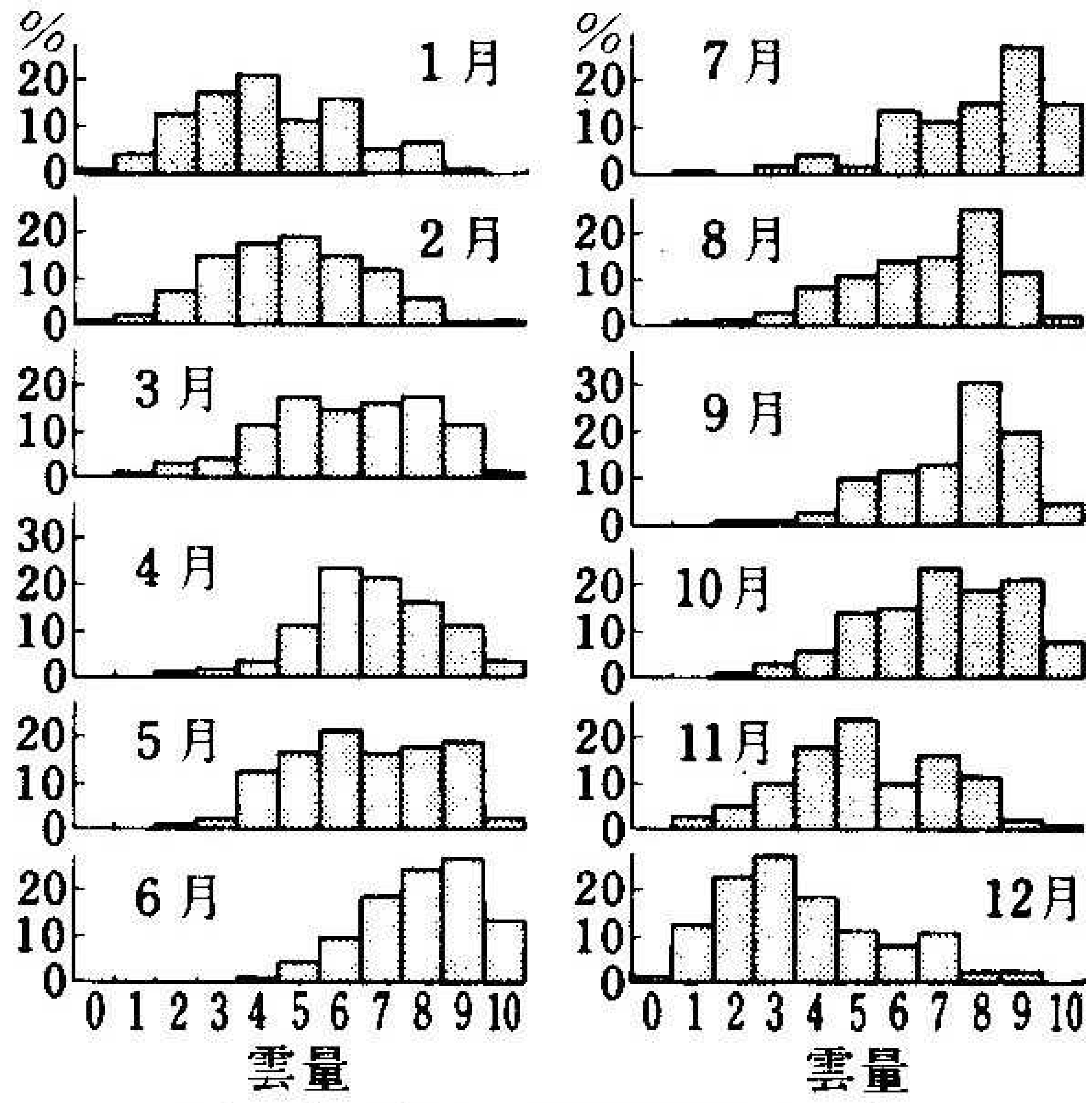


東京における6時の雲量のヒストグラム

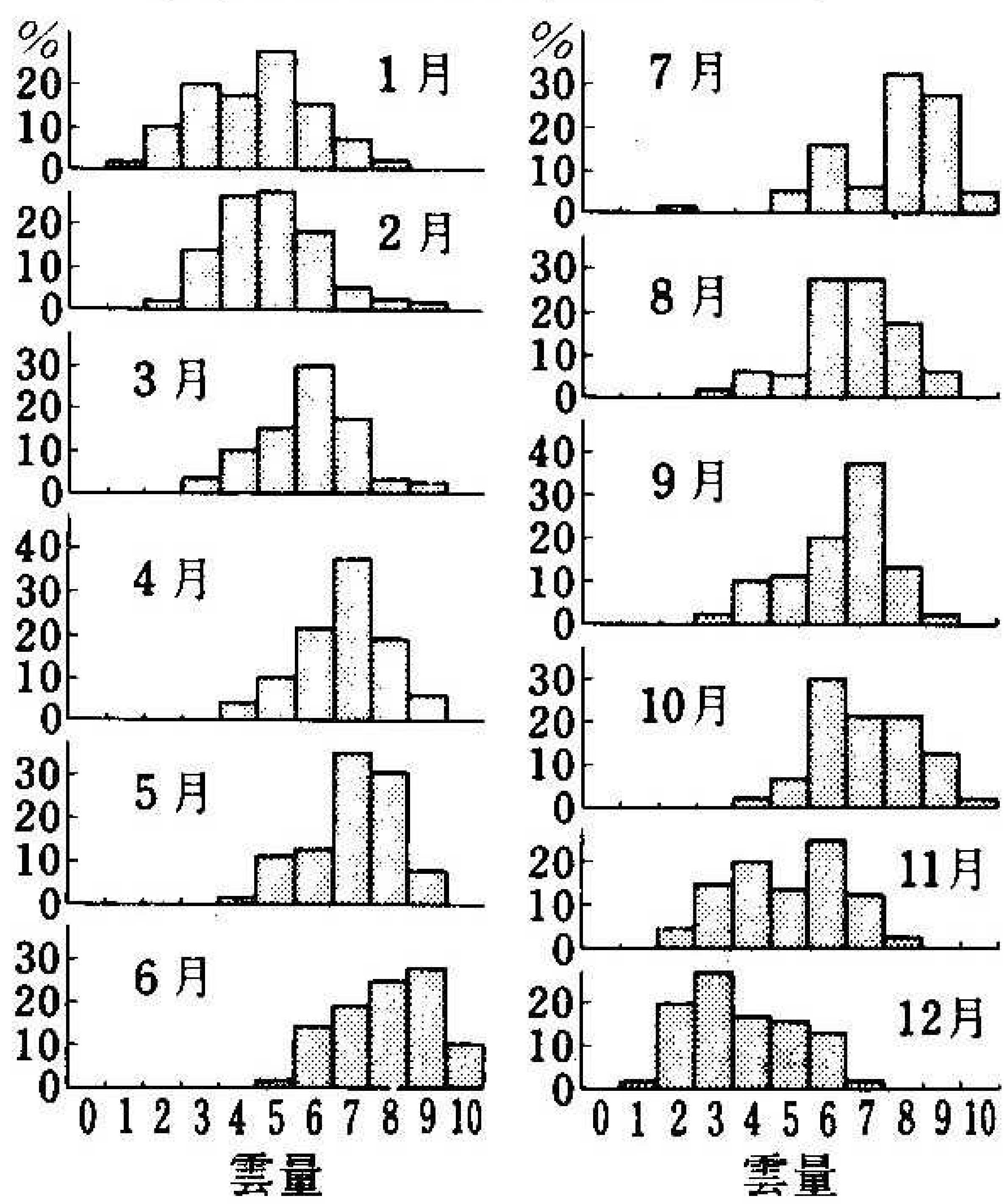
上のように両端で多く，中央で少ないU字型分布はあまり多くはありません。それで雲量のヒストグラムはよく引用されることになります。なぜ，このような分布をするのか，ということは地学（くわしくいえば，その中の気象学）の問題ですが，日平均雲量となると下のように右上りまたは左上りとなるし，さらに半旬平均になるとほぼ中央が多く，両端で低くなるというわけです。年平均雲量となるともっとその傾向がはっきり出てきます。



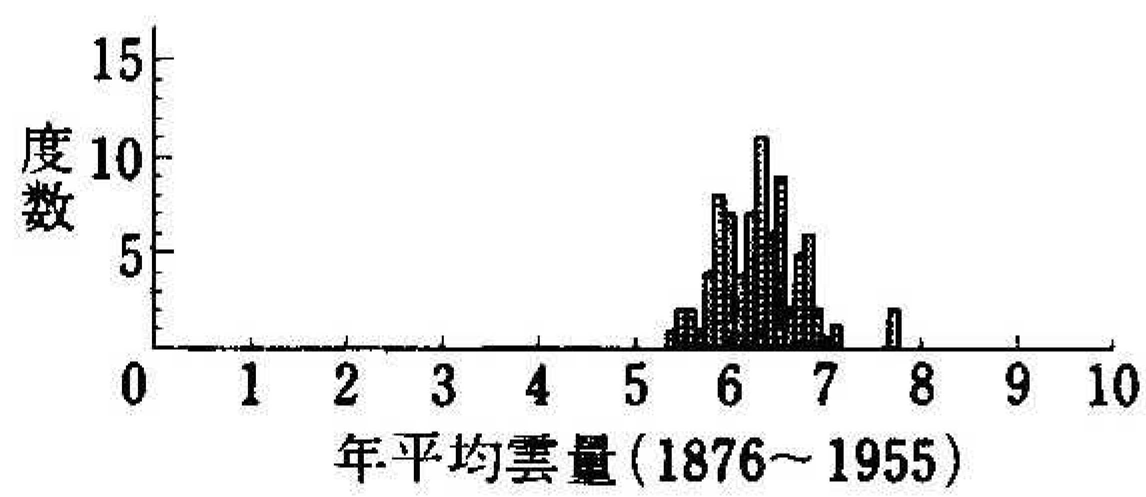
東京における日平均雲量の月別度数百分率の分布を示したもの（1891～1910）



半旬平均雲量の月別度数百分率の分布を示したもの（1891～1910）



旬平均雲量の月別度数百分率の分布を示したもの



東京における年平均雲量の度数分布（1876～1955）。最大値は7.8，最大値は5.4で，最も度数の多いのは6.3である。

＊　＊　＊

◆ ともあれ，こうしてみると，ヒストグラムの直観性を否定するわけにはいきません。しかし，これから，他の代表値を求めるにはむかない，という欠点はあるのです。

1回目　年　月　日
2回目　年　月　日
3回目　年　月　日

○平均値・中央値・最頻値とは

◆度数分布でわかったことを，もっと直截簡明に表すのが代表値というものです。そこに平均値や中央値や，が出てくる。

◆ 資料を大きさの順に並べたとき，その中央の値を**中央値**（メジアン）といいます。では，まず，これを：――

■練習 1. あるクラスの男子19人の身長を調べたところ，次のようであった。(単位cm)
中央値 M_e を求めよ。

171.1 164.7 172.6 177.8 172.9 170.5
170.7 177.2 156.0 162.9 169.7 169.8
165.0 183.5 165.3 152.8 162.2 178.4
172.0

ヒント　小さい方から順に並べてみますと右の表のようになります。したがって中央の値，つまり10位のものは170.5cm，これが中央値です。

順位	変量
1	152.8
2	156.0
3	162.2
4	162.9
5	164.7
6	165.0
7	165.3
8	169.7
9	169.8
10	170.5
11	170.7
12	171.1
13	172.0
14	172.6
15	172.9
16	177.2
17	177.8
18	178.4
19	183.5

なお，この場合は19人でしたから，ちょうど中央のものがありました。しかし，最後の第19位がなかったらどうするか。そのときは，中央の2つ第9位，第10位の平均をとるのです。

すなわち，

$$M_e=\frac{1}{2}(169.8+170.5)$$
$$=170.15$$

とするわけです。さらに，階級で与えられるときはどうするか，それが次の練習です。

■練習 2. 赤，白2種類の顆粒（かりゅう）（小さな粒）を入れたカプセル100個について，カプセル中の赤顆粒の数を調査したところ，下表の結果を得た。1カプセル中の赤顆粒数の中央値 M_e を求めよ。　　（静岡薬大）

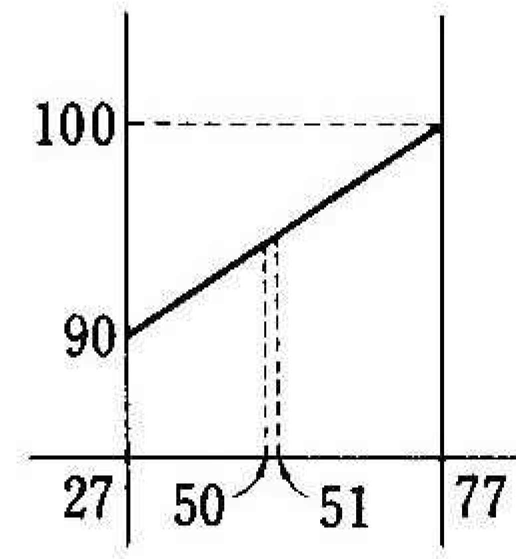
1カプセル中の赤顆粒数	カプセル数
71— 80	7
81— 90	20
91—100	50
101—110	22
111—120	1

ヒント　第27位が90，第77位が100と考えて比例するとしますと，第50位は

$$90+\frac{50-27}{77-27}\times10$$

第51位は

$$90+\frac{51-27}{77-27}\times10$$



とすることができましょう。その平均をとればいいでしょう。

それは

$$M_e=\frac{1}{2}\left(\frac{50-27}{50}\times10+\frac{51-27}{50}\times10\right)+90$$
$$=94.7$$

となります。　　　　答　94.7

*　　*　　*

◆ 次は**最頻値**（サイヒンチ）（モード）です。これは，度数の最も大きい階級の階級値をいいます。ではこれを：――

■練習 3. 次の度数分布表から最頻値 M_o を求めよ。

階級値(x_i)	10	12	14	16	18	20	計
度　数(f_i)	4	5	7	3	2	1	22

ヒント　度数のもっとも多いのはいうまでもなく14ですから M_o は14です。

*　　*　　*

◆ **平均値**とは，いうまでもなく，変量が x_1, x_2, ……, x_n のとき

$$\overline{x}=\frac{1}{n}\sum_{k=1}^{n}x_k \qquad \cdots\cdots(*)$$

をいいます。

■練習 4. 右の表は東京地方の大地震を示したものである。その間隔の平均を求めよ。

年代	間隔
818	
878	60
1257	379
1293	36
1495	202
1605	110
1649	44
1703	54
1812	109
1855	43
1894	39
1923	29

ヒント いうまでもなく

$$\frac{1}{11}\{60+379+36+202+110+44+54+109+43+39+29\}$$

$$=\frac{1}{11}\times1105=100.5$$

これによるとほぼ100年に1回大地震の起こっていることがわかります。

しかも，100年以下のものは7個，つまり64%におよんでいます。

■練習 5. あるクラス40人の生徒の体重を測って次のような度数分布表を得た。これから体重の平均値$\overline{x}$を小数第1位まで求めよ。 (東北学院大)

体重(kg)	40	41	42	43	44	45	46	47	48	49
人数(人)	1	0	2	5	7	10	8	4	1	2

ヒント もちろん上の(*)によって求めるわけですが，階級値 x_1, x_2, ……, x_n, その度数が f_1, f_2, ……, f_n のとき平均値は

$$\overline{x}=\frac{1}{N}\sum_{k=1}^{n}x_kf_k \quad \left(N=\sum_{k=1}^{n}f_k\right)$$

で与えられます。そこで，次のように計算すればよいでしょう。

体重	40	41	42	43	44	45	46	47	48	49	計
人数	1	0	2	5	7	10	8	4	1	2	40
積	40	0	84	215	308	450	368	188	48	98	1799

5

$$\therefore \quad \overline{x}=\frac{1799}{40}=44.9.7$$

答 45

* * *

◆ 度数のバラツキが大きくないとき，**仮平均**を使うと計算がグッと楽になることがあります。つまり，仮平均をx_0，階級の幅を c，とするとき，

$$u_k=\frac{x_k-x_0}{c}$$

とおくと

$$\overline{x}=x_0+c\overline{u}$$

となります。では，その例を：——

■練習 6. 得点x_kをとった人数をfとすると次の度数分布表を得た。平均値を求めよ。

x_k	85	95	105	115	125	135	145	計
f	4	3	8	12	7	4	2	40

ヒント x_kの中央の値を仮平均にとってみましょう。つまり $x_0=115$, その上で次のような表をつくってみましょう。もちろん $c=10$ です。

$$u_k=\frac{x_k-x_0}{c}$$

を求めると右の表のようになって，

x_k	f	u_k	u_kf
85	4	−3	−12
95	3	−2	−6
105	8	−1	−8
115	12	0	0
125	7	1	7
135	4	2	8
145	2	3	6
計	40		−5

$$\overline{u}=\frac{-5}{40}$$

$$\therefore \quad \overline{x}=115+10\overline{u}$$

$$=115+10\cdot\frac{-5}{40}$$

$$=113.8$$

答 113.8

* * *

◆ このように，単に平均をとるというだけでも，いろいろなことがあるのです。そして，それは，必ずやってみないとわからないことがあるものです。

なお，ついでながら，ここでいう平均値はいわゆる相加平均あるいは算術平均といわれるものですが，対象によっては相乗平均（幾何平均）や調和平均をとらなければならないものもあるのです。

資料の標準偏差と分散

1回目　年　月　日
2回目　年　月　日
3回目　年　月　日

◆平均値を代表者とするとき，代表者の手に負えないもの，それは，全体のバラツキの度合です。

◆ 変量のバラツキの度合を表すのが**分散**と**標準偏差**です。ところで，変量 x_k の分散 σ^2 は，その平均を $\overline{x}$ とするとき

$$\sigma^2=\frac{1}{N}\sum_{k=1}^{n}(x_k-\overline{x})^2f_k \quad \cdots\cdots(*)$$

$$=\frac{1}{N}\sum_{k=1}^{n}x^2{}_kf_k-\overline{x}^2 \quad \cdots\cdots(**)$$

で与えられます。ただし，$N=\sum_{k=1}^{n}f_k$ です。

＊　＊　＊

では，何はともあれ，次の練習をやってみませんか。

■練習 1. 変量 1，2，3，4，5 の分散を求めよ。

ヒント　公式(*)によって次の表のように計算してみましょう。

x_k	1	2	3	4	5	和
f_k	1	1	1	1	1	5
x_kf_k	1	2	3	4	5	15
$x_k-\overline{x}$	−2	−1	0	1	2	／
$(x_k-\overline{x})^2$	4	1	0	1	4	10

少しくどいがガマンして公式通りやってみますと第3段から

$$\overline{x}=\frac{15}{5}=3$$

これを使って第4段，したがって，第5段が求められ，よって

$$\sigma^2=\frac{10}{5}=2$$

となります。　　　答　$\sigma^2=2$

■練習 2. 資料 x_1，x_2，……，x_{50} の平均値が 9 で，平方和が4900のとき，分散を求めよ。

ヒント　上の公式(**)において

$$\overline{x}=9,\ \sum_{k=1}^{50}x_k{}^2=4900,\ N=50$$

$f_k=1$ ($k=1$, 2, …, 50) というのですから，

$$\sigma^2=\frac{1}{50}\sum_{k=1}^{50}x_k{}^2\cdot1-\overline{x}^2$$

$$=\frac{1}{50}\times4900-9^2$$

$$=98-81=17$$

答　17

＊　＊　＊

◆ では次に度数分布から求める問題をやってみませんか。

■練習 3. 右の度数分布表は高校1年男子44人の座高について調べた結果である。これについて，分散と標準偏差を求めよ。

座　高(cm) (以上)(未満)	人数 人
82.5～84.5	5
84.5～86.5	9
86.5～88.5	11
88.5～90.5	11
90.5～92.5	6
92.5～94.5	2
計	44

ヒント　まず階級から階級値を求め，それについて下の表のように計算すればよいでしょう。なお，階級値は階級の中央の値をとります。

階級値 (x)	度数 (f)	xf	x^2	x^2f
83.5	5	417.5	6972.25	34861.25
85.5	9	769.5	7310.25	65792.25
87.5	11	962.5	7656.25	84218.75
89.5	11	984.5	8010.25	88112.75
91.5	6	549.0	8372.25	50233.50
93.5	2	187.0	8742.25	17484.50
計	44	3870.0	／	340703.00

$$\overline{x}=\frac{3870.0}{44}\fallingdotseq87.95\fallingdotseq88.0$$

$$\therefore\quad \sigma^2=\frac{1}{44}\times340703.00-87.95^2$$

$=7743.25-7735.20$

$\fallingdotseq 8.0$

$$\sigma \fallingdotseq 2.83$$

このように数が大きくなると計算がたいへんです。そこで仮階級値，仮平均を使って求める方がよいのです。それは

$$\boldsymbol{\sigma^2=c^2\left(\frac{1}{N}\sum_{i=1}^{n}u_i^2f_i-\overline{u}^2\right)}$$

です。ここに，cは階級の幅で，仮平均をx_0とするとき

$$\boldsymbol{u_i=\frac{x_i-x_0}{c},\ \ \overline{u}=\frac{1}{N}\sum_{i=1}^{n}u_if_i}$$

です。ここでは $x_0=87.5$，$c=2$ としますと

$$u_i=\frac{x_i-87.5}{2}$$

で，練習 3. は下のようになります。

x	u	f	uf	u^2f
83.5	−2	5	−10	20
85.5	−1	9	−9	9
87.5	0	11	0	0
89.5	1	11	11	11
91.5	2	6	12	24
93.5	3	2	6	18
計		44	10	82

$$\overline{u}=\frac{10}{44}\fallingdotseq 0.2273$$

$$\frac{1}{N}\sum u^2f=\frac{82}{44}\fallingdotseq 1.8636$$

$$\therefore\quad \sigma^2=2^2(1.8636-0.2273^2)$$

$$\fallingdotseq 7.25$$

$$\sigma=\sqrt{7.25}=2.69$$

こうしてみると，仮平均の方がいかに簡単で，いかに正確であるかがわかるでしょう。

＊　＊　＊

◆ では，やや，総合的な問題をやってみませんか。

■練習 4. 5 個の正数がある。おのおのの 2 乗の和が 151 で，相異なる 2 つの数の積の和が 237 である。このとき，この 5 個の正数の平均値と分散を求めよ。　（東京医大）

ヒント　平均値をmとしますと

$$m^2=\left(\frac{1}{5}\sum_{i=1}^{5}x_i\right)^2=\frac{1}{25}\left(\sum_{i=1}^{5}x_i^2+2\sum_{i}\sum_{j}x_ix_j\right)$$

$$=\frac{1}{25}(151+2\times 237)=25$$

ところが $x_i>0$ ですから $m>0$

$$\therefore\quad m=5$$

次に分散を s^2 としますと

$$s^2=\frac{1}{5}\sum_{i=1}^{5}x_i^2-m^2=\frac{1}{5}\times 151-5^2$$

$$=\frac{26}{5}=5.2$$

答　平均値 5，分散 5.2

■練習 5. あることがらについて，n 人中賛成者の割合がpで，残りは反対者であるという。賛成者に 1，反対者に 0 という値Xを与えるとき，

(1)　変数Xの度数分布表を作成せよ。

(2)　変数Xの平均値mと，標準偏差σを求めよ。

ヒント　(1)　n 人中賛成者の割合がpだというのですからその人数は np 人でしょう。反対者はもちろん

$$(n-np)\text{人}=n(1-p)\text{人}$$

です。したがって，度数分布表は

階級値x	度数f
0	$n(1-p)$
1	np
計	n

(2)　上の表から，平均値mは

$$m=\frac{1}{n}\{0\times n(1-p)+1\times np\}$$

$$=p$$

また，分散 σ^2 は

$$\sigma^2=\frac{1}{N}\sum x^2-m^2$$

$$=\frac{1}{n}\{0^2\times n(1-p)+1^2\times np\}-p^2$$

$$=p(1-p)$$

$$\therefore\quad \sigma=\sqrt{p(1-p)}$$

統計の合併

1回目　年　月　日
2回目　年　月　日
3回目　年　月　日

◆いくつかのべつべつに行なわれた統計をまとめたいこともある。それが統計の合併というものです。

◆ べつべつに行った統計をまとめたらどうなるか，それがここの目的です。では，これをやってみませんか。

練習1. 右の表はA組，B組の同種の調査資料の度数分布表である。A組の平均値を M_1，B組の平均値を M_2 とするとき，2組を1つにまとめた総度数 n_1+n_2 の資料の平均値 M は

	A組	B組
変量	度数	度数
a_1	f_1	f_1'
a_2	f_2	f_2'
$\vdots$	$\vdots$	$\vdots$
a_n	f_n	f_n'
計	n_1	n_2

$$M=\frac{n_1M_1+n_2M_2}{n_1+n_2} \qquad \cdots\cdots(*)$$

で求められることを示せ。

ヒント　平均値の定義から

$$M_1=\frac{\sum a_i f_i}{n_1},\ M_2=\frac{\sum a_i f_i'}{n_2}$$

ところが

$$M=\frac{\sum a_i(f_i+f_i')}{n_1+n_2}$$

$$=\frac{\sum a_i f_i+\sum a_i f_i'}{n_1+n_2}$$

$$=\frac{n_1M_1+n_2M_2}{n_1+n_2} \qquad \text{Q.E.D.}$$

練習2. 右の表は，200人の生徒を3つのグループA，B，Cに分けて行った学力検査の結果をまとめたものである。

	人数	平均値	標準偏差
A	62	63.0	8.2
B	70	68.3	9.1
C	68	63.5	8.6
計	200		

生徒全体の平均値と標準偏差を小数第1位まで求めよ。　(大阪教育大)

ヒント　上の公式(*)をオボエテいなくとも平均値は求められるハズです。そして，それは：——

$$\bar{x}=\frac{1}{200}(62\times63.0+70\times68.3+68\times63.5)$$

$$=65.0$$

となります。そこで，A，B，C各グループの各人の得点を a_i，b_i，c_i としますと，標準偏差の定義から

$$(8.2)^2=\frac{1}{62}\sum a_i^2-(63.0)^2$$

$$(9.1)^2=\frac{1}{70}\sum b_i^2-(68.3)^2$$

$$(8.6)^2=\frac{1}{68}\sum c_i^2-(63.5)^2$$

ですから，

$$\sigma^2=\frac{1}{200}(\sum a_i^2+\sum b_i^2+\sum c_i^2)-(65.0)^2$$

$$=84.0408$$

$\therefore\ s\fallingdotseq9.2$　　　答 9.2

*　*　*

◆ では，次には，やや総合的な問題をやってみることにしましょう。

練習3. 男子 m 人，女子 n 人のクラスがある。男子の身長の平均は h_1，分散は s_1^2 であり，女子の身長の平均は h_2，分散は s_2^2 である。

(1) クラス全体の身長の平均 h と分散 s^2 を，m，n，h_1，h_2，s_1，s_2 を用いて表せ。

(2) 不等式 $s^2\geqq\dfrac{ms_1^2+ns_2^2}{m+n}$ が成り立つことを示せ。　(筑波大)

ヒント　(1)の前半は練習1.と同じですから，今

更やる必要もないでしょう。それは，

$$h=\frac{mh_1+nh_2}{m+n}$$

次に，分散も練習2.と本質的に同じですが，次にやってみましょう。ここで x_i, y_i は，それぞれ男子および女子各人の身長です。

$$s_1{}^2=\frac{1}{m}\sum x_i{}^2-h_1{}^2,\quad s_2{}^2=\frac{1}{n}\sum y_i{}^2-h_2{}^2$$

ですから

$$\sum x_i{}^2=m(s_1{}^2+h_1{}^2),\quad \sum y_i{}^2=n(s_2{}^2+h_2{}^2)$$

となります。そして，

$$s^2=\frac{1}{m+n}\left(\sum x_i{}^2+\sum y_i{}^2\right)-h^2$$

$$=\frac{m(s_1{}^2+h_1{}^2)+n(s_2{}^2+h_2{}^2)}{m+n}-\frac{(mh_1+nh_2)^2}{(m+n)^2}$$

$$=\frac{ms_1{}^2+ns_2{}^2}{m+n}$$

$$+\frac{(m+n)(mh_1{}^2+nh_2{}^2)-(mh_1+nh_2)^2}{(m+n)^2}$$

$$=\frac{ms_1{}^2+ns_2{}^2}{m+n}+\frac{mn}{(m+n)^2}(h_1-h_2)^2$$

(2)　$s^2-\dfrac{ms_1{}^2+ns_2{}^2}{m+n}$

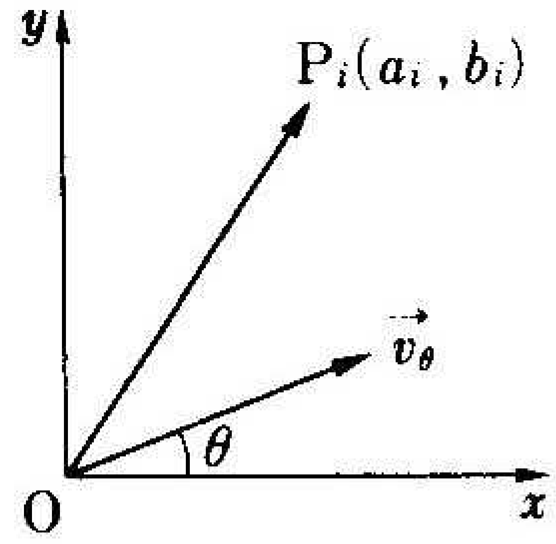
$$=\frac{mn}{(m+n)^2}(h_1-h_2)^2\geqq 0$$

等号は $h_1=h_2$ のとき成り立つのです。では，もうひとつ：——

練習 4. n 人の人に2つの検査A，Bを行ったところ，各人はそれぞれ数値 a_1, a_2, ……, a_n ; b_1, b_2, ……, b_n を得た。第 i 番目の人の検査結果を平面上の点 $P_i(a_i, b_i)$ $(i=1, 2, \cdots\cdots, n)$ で表すことにし，原点Oから x 軸と θ の角 $(0\leqq\theta\leqq\pi)$ をなす長さ1のベクトル $\vec{v_\theta}$ を引く。ベクトル $\vec{v_\theta}$ と $\overrightarrow{\mathrm{OP}_i}$ の内積 $c_i=(\vec{v_\theta}, \overrightarrow{\mathrm{OP}_i})$ $(i=1, 2, \cdots\cdots, n)$ を第 i 番目の人の検査A，Bの総合数値とする。ただし，$\{a_i\}$, $\{b_i\}$ はともに平均値0，標準偏差1とする。

(1)　$\{c_i\}$ の平均値を求めよ。

(2)　$\{c_i\}$ の標準偏差が最大になるような θ を求めよ。　　(慶大)



ヒント　(1)　$\vec{v_\theta}=(\cos\theta, \sin\theta)$ ですから，

$$c_i=(\vec{v_\theta}, \overrightarrow{\mathrm{OP}_i})$$

$$=(\cos\theta, \sin\theta)\cdot(a_i, b_i)$$

$$=a_i\cos\theta+b_i\sin\theta$$

$$\therefore\quad \bar{c}=\frac{1}{n}\sum_{i=1}^{n}(a_i\cos\theta+b_i\sin\theta)$$

$$=\left(\frac{1}{n}\sum_{i=1}^{n}a_i\right)\cos\theta+\left(\frac{1}{n}\sum_{i=1}^{n}b_i\right)\sin\theta$$

$$=\bar{a}\cos\theta+\bar{b}\sin\theta$$

ところが $\bar{a}=0$, $\bar{b}=0$ というのですから

$$\bar{c}=0$$

さあ，これで(1)は意外と楽にできた。次は(2)です。

(2)　$\{c_i\}$ の標準偏差を s としますと，

$$s^2=\frac{1}{n}\sum_{i=1}^{n}c_i{}^2-\bar{c}^2$$

$$=\frac{1}{n}\sum_{i=1}^{n}(a_i\cos\theta+b_i\sin\theta)^2-0^2$$

$$=(\cos^2\theta)\left(\frac{1}{n}\sum_{i=1}^{n}a_i{}^2\right)+(\sin^2\theta)\left(\frac{1}{n}\sum_{i=1}^{n}b_i{}^2\right)$$

$$+(2\sin\theta\cos\theta)\left(\frac{1}{n}\sum_{i=1}^{n}a_ib_i\right)$$

ところが $\{a_i\}$, $\{b_i\}$ の標準偏差が1だというのですから

$$1=\frac{1}{n}\sum_{i=1}^{n}a_i{}^2-0^2$$

$$1=\frac{1}{n}\sum_{i=1}^{n}b_i{}^2-0^2$$

$$\therefore\quad s^2=\cos^2\theta+\sin^2\theta+(\sin 2\theta)\left(\frac{1}{n}\sum_{i=1}^{n}a_ib_i\right)$$

$$=1+(\sin 2\theta)\left(\frac{1}{n}\sum_{i=1}^{n}a_ib_i\right)$$

だから $\sum_{i=1}^{n}a_ib_i>0$ のときには $\theta=\dfrac{\pi}{4}$ のとき最大，$\sum_{i=1}^{n}a_ib_i<0$ のときには $\theta=\dfrac{3\pi}{4}$ のとき最大，そして $\sum_{i=1}^{n}a_ib_i=0$ のとき一定です。

*　　*　　*

◆　このように統計を合併することは，実際上しばしば必要になります。しかし，統計量の説明を欠くことが多いので困るのです。

資料の整理についてのいろいろな問題

1回目　年　月　日
2回目　年　月　日
3回目　年　月　日

◆資料の平均と標準偏差しかわからなくてもいろいろな性質を導くことができます。その最重要なのはチェビシェフの不等式!!

◆ 変量のバラツキを表すのに分散や標準偏差がよく用いられますが，**レンジ**（範囲ともいいます）というのもあります。これは変量の最大値と最小値の差です。

■練習 1. 次のデータのレンジを求めよ。

$$4,\ 7,\ 8,\ 3,\ 1,\ 4$$

ヒント　最大値は8で最小値は1ですから，その差は

$$8-1=7$$

です。これがレンジです。Rという記号で表すことがあります。

■練習 2. 変量uのレンジをRとするとき，$x=cu+m$（c，mは定数）のレンジを求めよ。

ヒント　uの最大値をu_1，最小値をu_2とすれば

$$R=u_1-u_2$$

そして，xの最大値，最小値はcの符号によってちがい，一方はcu_1+m，他方はcu_2+mですから，求めるレンジは

$$|(cu_1+m)-(cu_2+m)|$$
$$=|c|(u_1-u_2)=|c|R$$

であることがわかります。

＊　＊　＊

◆ 平均値について，次の関係があります。

■練習 3. ある変量について，階級の幅がlの度数分布表から求めた平均値を$\overline{x}$，真の平均値をmとすれば，$|m-\overline{x}|\leqq\dfrac{l}{2}$であることを示せ。

ヒント　まず問題の意味を知らなければなりません。具体的な例でいきましょう。

あるクラスの男子20人の身長を調べたところ次のようになりました。（単位はcm）

171.1	164.7	172.6	177.8	172.9
170.5	170.7	177.2	156.0	162.9
169.7	169.8	165.0	183.5	165.3
152.8	162.2	178.4	172.0	156.8

この平均をとってみますと

$$m=168.595$$

となります。一方，上のデータから下のような度数分布表を作りますと次のようになります。これから平均値を求めてみますと，

身　長	x	f
150～155	152.5	1
155～160	157.5	2
160～165	162.5	3
165～170	167.5	4
170～175	172.5	6
175～180	177.5	3
180～185	182.5	1
計		20

$$\overline{x}=168.75$$

となります。

だから，

$$|m-\overline{x}|$$
$$=|168.595-168.75|$$
$$=0.155$$

となります。

ところで，級間は

$$l=5$$

ですから

$$\frac{l}{2}=2.5$$

で，確かに

$$0.155<2.5$$

つまり

$$|m-\overline{x}|\leqq\frac{l}{2}$$

となるのです。これを一般的に証明せよ，というのが，この問題です。

では，その証明を：――

階級値を x_1, x_2, ……, x_n とし，これに対応する各階級の度数を f_1, f_2, ……, f_n としましょう。もちろん

$$\sum_{i=1}^{n} f_i = N$$

です。さて：——

i 番目の階級に属する f_i 個の変量を

x_{i1}, x_{i2}, ……, x_{if_i} $(i=1,2,\cdots,n)$

としますと，階級の幅が l ですから

$$|x_{i1}-x_i| \leqq \frac{l}{2},\ |x_{i2}-x_i| \leqq \frac{l}{2},\ \cdots\cdots$$

$$\cdots\cdots,\ |x_{if_i}-x_i| \leqq \frac{l}{2}$$

以上から

$$|(x_{i1}+x_{i2}+\cdots\cdots+x_{if_i})-x_if_i|$$
$$=|(x_{i1}-x_i)+(x_{i2}-x_i)+\cdots+(x_{if_i}-x_i)|$$
$$\leqq |x_{i1}-x_i|+|x_{i2}-x_i|+\cdots\cdots+|x_{if_i}-x_i|$$
$$\leqq \frac{l}{2}+\frac{l}{2}+\cdots\cdots+\frac{l}{2}=\frac{l}{2}\cdot f_i$$

$$\therefore\ |m-\overline{x}|=\left|\frac{1}{N}\sum_{k=1}^{N}x_k-\frac{1}{N}\sum_{i=1}^{n}x_if_i\right|$$
$$=\frac{1}{N}\left|\sum_{i=1}^{n}\{(x_{i1}+x_{i2}+\cdots\cdots+x_{if_i})-x_if_i\}\right|$$
$$\leqq \frac{1}{N}\cdot\sum_{i=1}^{n}\frac{l}{2}f_i=\frac{1}{N}\cdot\frac{l}{2}\sum_{i=1}^{n}f_i=\frac{1}{N}\cdot\frac{l}{2}N$$
$$=\frac{l}{2}$$

$$\therefore\ |m-\overline{x}| \leqq \frac{l}{2}$$

ヤレヤレ，以外トメンドウだったなあ。

* * *

◆ チェビシェフの不等式といわれるものがあります。それは：——

N 個の変量 x_1, x_2, ……, x_N の平均値を $\overline{x}$，標準偏差を s，また，$k>1$ とする。このN個の変量のうち

$$\overline{x}-ks \leqq x_i \leqq \overline{x}+ks$$

となるものの個数を n とすると

$$\frac{n}{N}>1-\frac{1}{k^2}$$

が成り立つ

〈証明〉 $\overline{x}-ks \leqq x_i \leqq \overline{x}+ks$

を変形すると

$$-ks \leqq x_i-\overline{x} \leqq ks$$

となります。さらにこれを変形すると

$$(x_i-\overline{x})^2 \leqq k^2s^2 \qquad \cdots\cdots(*)$$

となります。(*) を満足するものは n 個ありますから，残りの $N-n$ 個については

$$(x_i-\overline{x})^2 > k^2s^2$$

$$\therefore\ Ns^2=\sum_{i=1}^{N}(x_i-\overline{x})^2>(N-n)k^2s^2$$

両辺を Nk^2s^2 で割ると

$$\frac{1}{k^2}>\frac{N-n}{N}$$

$$\therefore\ \frac{n}{N}>1-\frac{1}{k^2}$$

Q.E.D.

では，その応用を：——

練習 4. 210人の高校1年生男子の胸囲を測ったら，平均値は 81.2cm，標準偏差は 4.22cm であるとき，76.2cm 以上，86.2 cm以下の人は何人以上いるといえるか。

ヒント チェビシェフの不等式において

$$N=210,\ \overline{x}=81.2,\ s=4.22$$

ですから

$$(76.2-\overline{x})^2=(86.2-\overline{x})^2$$
$$=5^2 \leqq k^2 \times 4.22^2$$

$$\therefore\ k^2 \geqq \frac{5^2}{4.22^2}$$

$$\therefore\ \frac{n}{210}>1-\frac{1}{k^2}>1-\frac{4.22^2}{5^2}$$

$$\fallingdotseq \frac{7.192}{25}$$

$$\therefore\ n>60.4$$

このことから，$n \geqq 61$

答 61人以上

* * *

◆ ここでは資料の性質についてとくに立入った性質がわかっていない場合について扱いました。しかし，実際には，資料の本来の性質からくるいろいろな制約があるものです。そのため，もっと立ち入った性質が求められます。それが「統計学」の中心となります。

◑ 確率変数とは何か

1回目　年　月　日
2回目　年　月　日
3回目　年　月　日

◆よく知っていることでも表現を変えるととまどうものだ。確率は知っている。変数は知っている。しかし，確率変数は，となると……

◆ 具体的な例からいきましょう。まず，次の練習を：——

■練習 1. 1つのさいころを振るとき，1，2，……，6の目の出る確率を求めよ。

ヒント いうまでもなく，いずれの目が出る確率も$\dfrac{1}{6}$です。このとき，

1，2，……，6 を確率変数

といい，各確率変数に対応する確率はいずれも$\dfrac{1}{6}$である，というのです。

■練習 2. 10本のくじがあって，そのうち2本が当たりくじとする。3本引いて当たったくじの本数をXとするとき，X本当たる確率p_Xを求めよ。

ヒント 当たりくじは2本しかないのですから$p_3=0$はいうまでもありません。次に，

$$p_2=\frac{{}_2C_2\cdot{}_8C_1}{{}_{10}C_3}=\frac{1}{15}$$

$$p_1=\frac{{}_2C_1\cdot{}_8C_2}{{}_{10}C_3}=\frac{7}{15}$$

$$p_0=\frac{{}_8C_3}{{}_{10}C_3}=\frac{7}{15}$$

となります。この結果を表にまとめてみますと，下のようになります。

X	0	1	2	3	計
p_X	$\frac{7}{15}$	$\frac{7}{15}$	$\frac{1}{15}$	0	1

つまり確率変数は0，1，2，3で，それに対応する確率はそれぞれ$\dfrac{7}{15}$，$\dfrac{7}{15}$，$\dfrac{1}{15}$，0というわけです。この確率の値の分布を **確率分布** といいます。では，もう1つ：——

■練習 3. 2枚の硬貨を同時に投げるとき，表の出る枚数Xを確率変数として，それに対応する確率p_Xを求めよ。

解 Xは0，1，2の値をとり，その確率は次のようである。

$$p_0=\frac{1}{4},\quad p_1=\frac{2}{4}=\frac{1}{2},\quad p_2=\frac{1}{4}$$

答 $p_0=\dfrac{1}{4}$，$p_1=\dfrac{1}{2}$，$p_2=\dfrac{1}{4}$

*　*　*

◆ 一般に確率変数Xが$x_k(k=1,\ 2,\ \cdots\cdots,\ n)$の値をとり，$x_k$に対応する確率を$p_k$とするとき，

$$\sum_{k=1}^{n}x_kp_k=x_1p_1+x_2p_2+\cdots\cdots+x_np_n$$

のことを$E(X)$や$\overline{X}$やmで表して，確率変数Xの **平均** または **期待値** といいます。

また，上の平均値mを用いて

$$\sum_{k=1}^{n}(x_k-m)^2p_k$$
$$=(x_1-m)^2p_1+(x_2-m)^2p_2+\cdots$$
$$\cdots+(x_n-m)^2p_n$$

を$E\{(X-m)^2\}$や$V(x)$やσ^2で表して，これを確率変数Xの **分散**（ぶんさん）といいます。なおσはΣの小文字で，もちろんシグマと読みます。いいかげんに書くと6と混同するからご注意!!

最後に分散の正の平方根を **標準偏差** といいます。すなわち

$$\sigma=\sigma(X)=\sqrt{V(X)}$$

で，σの代わりに小文字のsを使うこともあります。それは標準の頭文字です!!

*　*　*

◆ では，次にややめんどうな問題にいきましょう。

練習 4. 1と書いたカード1枚，2と書いたカード2枚，……，nと書いたカードn枚がある。この中から1枚のカードをとり出すとき，カードの示す数Xを確率変数とする。

(1) $X=k$ である確率 p_k を求めよ。

(2) Xの平均値 m を求めよ。

(3) Xの標準偏差 σ を求めよ。

（札幌医大）

ヒント (1) カードの数は

$$1+2+3+\cdots\cdots+n=\frac{1}{2}n(n+1)$$

で，kと書いたカードがk枚あるのですから

$$p_k=\frac{k}{\frac{1}{2}n(n+1)}=\frac{2k}{n(n+1)}$$

です。

(2) 平均値は期待値に等しいのですから

$$1\cdot\frac{2\cdot 1}{n(n+1)}+2\cdot\frac{2\cdot 2}{n(n+1)}+\cdots$$

$$\cdots\cdots+n\cdot\frac{2\cdot n}{n(n+1)}$$

$$=\frac{2}{n(n+1)}(1^2+2^2+\cdots\cdots+n^2)$$

$$=\frac{2}{n(n+1)}\cdot\frac{1}{6}n(n+1)(2n+1)=\frac{2n+1}{3}$$

となりましょう。

(3) 標準偏差σに対して

$$\sigma^2=\sum_{k=1}^{n}\left(k-\frac{2n+1}{3}\right)^2 p_k$$

$$=\sum_{k=1}^{n}\left(k^2-2\,\frac{2n+1}{3}\cdot k+\frac{(2n+1)^2}{9}\right)\frac{2k}{n(n+1)}$$

$$=\frac{2}{n(n+1)}\sum_{k=1}^{n}\left(k^3-2\,\frac{2n+1}{3}k^2+\frac{(2n+1)^2}{9}k\right)$$

$$=\frac{2}{n(n+1)}\left\{\frac{1}{4}n^2(n+1)^2\right.$$

$$-2\,\frac{2n+1}{3}\cdot\frac{1}{6}n(n+1)(2n+1)$$

$$\left.+\frac{(2n+1)^2}{9}\cdot\frac{1}{2}n(n+1)\right\}$$

$$=\frac{(n+2)(n-1)}{18}$$

$$\therefore\quad \sigma=\frac{\sqrt{2(n+2)(n-1)}}{6}$$

*　　*　　*

◆ では，もう1問やってみましょう。

練習 5. n個の値 $x_i(i=1,\ 2,\ \cdots\cdots,\ n)$の平均（相加平均）および標準偏差をそれぞれ M_x，S_x とし，n個の値 $y_i=ax_i+b$ $(i=1,\ 2,\ ,\ \cdots\cdots,\ n)$ の平均（相加平均）および標準偏差をそれぞれ M_y，S_y とするとき，M_x と M_y との間，および S_x と S_y との間に成り立つ関係式を求めよ。ただし，aおよびbは定数とする。

（福島県医大）

ヒント まず，平均については

$$M_y=\frac{(ax_1+b)+(ax_2+b)+\ \cdots\ +(ax_n+b)}{n}$$

$$=a\frac{x_1+x_2+\cdots\cdots+x_n}{n}+b$$

$$=aM_x+b$$

次に，定義から

$$S_y=\sqrt{\frac{1}{n}\sum_{k=1}^{n}\{(ax_k+b)-(aM_x+b)\}^2}$$

$$=\sqrt{\frac{1}{n}\sum_{k=1}^{n}a^2(x_k-M_x)^2}$$

$$=|a|\sqrt{\frac{1}{n}\sum_{k=1}^{n}(x_k-M_x)^2}$$

$$=|a|S_x$$

答 $M_y=aM_x+b$，$S_y=|a|S_x$

注 $\sqrt{a^2}=|a|$ にご注意!!

*　　*　　*

◆ それにしても，確率の問題は，何かピンとこないものが残るでしょう。それもそのハズ。確率を数学として扱ったのはパスカルですが，その根本についてはパスカル，デカルト，スピノザなどによってくり返しとりあげられたのでした。しかし，今でもまだスッキリしたとはいえません。まして，これを勉強する人が，そのために悩むのも当然というものでありましょう。

○確率分布とは何か

1回目　年　月　日
2回目　年　月　日
3回目　年　月　日

◆確率分布とは何か，というとめんどうでも，
確率分布表を作れ，といえばやさしくなる。
人間とは何かより人間がどれかのほうが簡単。

◆ 確率変数の項（☞p.128）に軽く目を通してから，このセクションをみてください。

確率変数Xが x_1, x_2, ……, x_n の値をとり，そのときの確率がそれぞれ p_1, p_2, ……, p_n であるとき，次のように表に表すことができます。これを **確率分布表** といいます。

X	x_1	x_2	x_3	……	x_n	計
$p(X)$	p_1	p_2	p_3	……	p_n	1

そして，この表に示されたような $p(X)$ の分布が **確率分布** というわけです。いや，こんな定義よりも具体的な練習をやってみるほうがいいでしょう。

＊　＊　＊

◆ まず，さいころの問題です。

練習 1. 2つのさいころを振ったとき，出る目の数の和をXとし，それに対応する確率を $p(X)$ とするとき，$p(X)$ の確率分布表を作れ。

（解） 2つのさいころの目の和は 2, 3, …… 12という値をとり，その場合の数は下の表のようになる。

目の和	2	3	4	5	6	7	8	9	10	11	12	計
場合の数	1	2	3	4	5	6	5	4	3	2	1	36

全体の場合の数は $6\times6=36$ であるから，求める確率分布表は下のようになる。

X	2	3	4	5	6	7	8	9	10	11	12	計
$p(X)$	$\frac{1}{36}$	$\frac{1}{18}$	$\frac{1}{12}$	$\frac{1}{9}$	$\frac{5}{36}$	$\frac{1}{6}$	$\frac{5}{36}$	$\frac{1}{9}$	$\frac{1}{12}$	$\frac{1}{18}$	$\frac{1}{36}$	1

（注） 上の $p(X)$ の値は約さないほうが見やすいから約さない人もある。そこは趣味（ないし，実用上）の問題である。

練習 2. 白球4個，黒球が5個入っている袋の中から5個とり出すとき，黒球の含まれる個数の確率分布を求めよ。

（ヒント） 全部で9個，この中から5個とり出す場合の数は ${}_9C_5$ です。

そして，

黒球が含まれない場合の数は明らかに0

黒球が1個含まれている場合の数は

$${}_4C_4\cdot{}_5C_1=5$$

黒球が2個含まれている場合の数は

$${}_4C_3\cdot{}_5C_2=40$$

といったぐあい。

結果は次の通りです。

答

X	0	1	2	3	4	5	計
$p(X)$	0	$\frac{5}{126}$	$\frac{40}{126}$	$\frac{60}{126}$	$\frac{20}{126}$	$\frac{1}{126}$	1

＊　＊　＊

◆ 今まで練習してきたのは，すべて，確率変数のとる値が不連続でした。このようなとき，Xを **離散形の確率変数** といいます。それに対し，変数Xが連続な値をとることもあります。このとき，X を **連続形の確率変数** といい，$X<x$ である確率 $P(X<x)$ を X の **分布関数** といいます。そして，X の分布関数の導関数を，**確率密度関数** というのです。つまり

$$F(x)=P(X<x)$$

としますと，確率密度関数 $f(x)$ は

$$f(x)=F'(x)$$

で与えられます。こんなことばは覚える必要はありません。

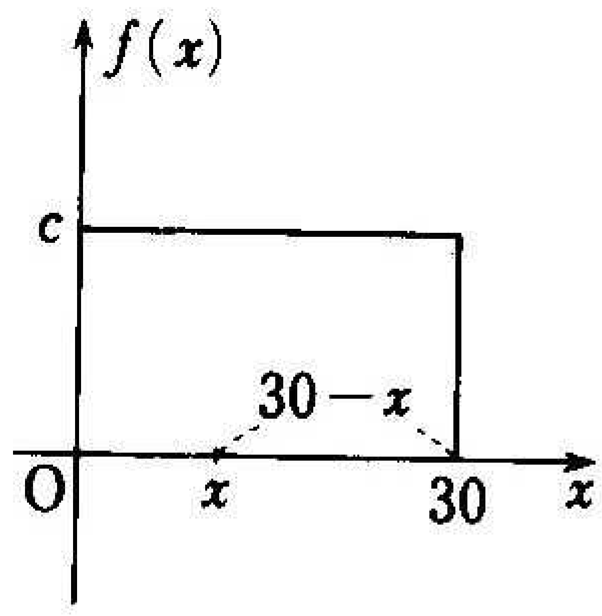
では，次の練習をやってみませんか。

■練習 3. あるバスの停留所では，毎時 0 分，30分の 2 回発車する。この発車時刻を知らないで，ここに来る人の待たされる時間の平均値（期待値）を求めよ。

(ヒント) ある30分間，0 分〜30分について考えればいいでしょう。

そうすると，停留所に来る時刻 x 分は

$$0<x\leqq 30$$



の値をとる変数で，これが確率変数です。そして，この値をまったく同様な確からしさでとるのですから，確率密度 $f(x)$ は一定です。つまり

$$f(x)=\text{const}$$

これを c とすると

$$\int_0^{30} c\,dx=1 \quad \therefore \quad 30c=1 \quad \therefore \quad c=\frac{1}{30}$$

かくして待ち時間の期待値 E は

$$E=\int_0^{30}\frac{1}{30}(30-x)\,dx=\frac{1}{30}\left[30x-\frac{x^2}{2}\right]_0^{30}$$

$$=\frac{1}{30}\left\{30\cdot 30-\frac{30^2}{2}\right\}$$

$$=15\ (分)$$

である。（思エバアタリマエ!!）

(答) 15分

■練習 4. 連続的な値をとる確率変数 x があって，その確率密度が A を定数として

$$f(x)=\begin{cases} Ax^2+1 & \left(0\leqq x\leqq \dfrac{1}{2}\right) \\ 0 & \left(x<0,\ \ x>\dfrac{1}{2}\right) \end{cases}$$

とするとき，$P(x\leqq a)=0.232$ となるような a の値を求めよ。　（長崎大）

(ヒント) まず，何はともあれ，A をきめよう。

$$\int_0^{\frac{1}{2}}(Ax^2+1)\,dx=\left[\frac{Ax^3}{3}+x\right]_0^{\frac{1}{2}}$$

$$=\frac{A}{24}+\frac{1}{2}=1$$

ですから，

$$A=12$$

そこで，

$$P(x\leqq a)=\int_0^a(12x^2+1)\,dx$$

$$=\left[4x^3+x\right]_0^a=4a^3+a$$

$$\therefore \quad 4a^3+a=0.232 \qquad \cdots\cdots(*)$$

ところで，これを解くのが問題だ!!

変形すると

$$1000a^3+225a-58=0$$

1000の約数で58の約数を割ったものをすべて入れてみればよいハズだが，それはたいへんだ。もとに戻って，(*) の両辺のグラフをかいて交点の座標を求めてみたらどうだろう。心は千々に乱れるところである!!

しかし，a が 1 より小さいことは確かだから a^3 は a に比べてかなり小さいにちがいない。してみると，a は 0.232 の近くの値なのであろう。

そうだ，ともあれ，0.2 より大か小かを調べてみようか？

そこで，これを代入してみますと，

$$4\times(0.2)^3+0.2=0.232$$

オヤオヤ，コレハ驚イタ。ちょうど 0.232 になったではないか。

(答) 0.2

(注) それにしても，この問題はちょっと気になるね。0.2 を入れてみたい気分にならない限り，たいへんだったにちがいないからだ。問題を作るほうの立場と，解かされる立場のちがいを，こんなに見せつけた問題も少ない。

■練習 5. X の確率密度関数 $f(x)$ が

$$f(x)=\begin{cases} Ax^2 & (0\leqq x\leqq 1) \\ 0 & (x<0,\ 1<x) \end{cases}$$

で与えられるとき，A の値を求めよ。また，$P\left(x\leqq\dfrac{1}{2}\right)$ を求めよ。

(ヒント) $\displaystyle\int_0^1 Ax^2\,dx=1 \quad \therefore \quad A=3$

したがって，

$$P\left(x\leqq\frac{1}{2}\right)=\int_0^{\frac{1}{2}}3x^2\,dx=\frac{1}{8}$$

○（確率変数の）平均とは何か

1回目　年　月　日
2回目　年　月　日
3回目　年　月　日

◆平均というコトバの妖しさ。確率変数の平均にいたって，平均はその本領を発揮するとみえた。それ，是か非か。

◆ 変量Xの度数分布表が次のようであったとしましょう。

変量Xの値	x_1　x_2　x_3 ……　x_n
度　　数	f_1　f_2　f_3 ……　f_n

このとき

$$\bar{x}=\frac{f_1x_1+f_2x_2+\cdots\cdots+f_nx_n}{N}$$

を **変量Xの平均** といいます。ここに，Nは

$$N=f_1+f_2+\cdots\cdots+f_n$$

です。なお，平均を表すには$\bar{x}$もよく使われますが $M(x)$ も使います。このMは平均(Mean) の頭文字です。

＊　＊　＊

◆ では，さっそく具体的な例にいきましょう。

練習 1. ある体格検査で右のような身長の度数分布表が得られた。これから平均の身長を求めよ。

（東大）

身長(cm)	人数
150—155	3
155—160	6
160—165	32
165—170	17
170—175	10
175—180	2

（解） 160 を仮平均にとって

$$(152.5-160)\times3+(157.5-160)\times6$$
$$+(162.5-160)\times32+\cdots\cdots$$
$$=(-7.5)\times3+(-2.5)\times6+(2.5)\times32$$
$$+7.5\times17+12.5\times10+17.5\times2$$
$$=330$$

ゆえに，求める平均は

$$160+\frac{330}{70}=164.7$$

答　164.7cm

＊　＊　＊

◆ 確率変数Xがn個の値 $x_1, x_2, \cdots\cdots, x_n$ のいずれかの値を必ずとり，その値をとる確率がそれぞれ p_1, p_2, ……, p_n であるとき

$$m=\sum_{i=1}^{n}p_ix_i$$

をXの **平均値** または **期待値** といいます。では，これをやってみませんか。

練習 2. 3面に数字1を刻み，他の3面にそれぞれ2, 3, 4の数字を刻んださいころがある。このようなさいころ2個を同時に振って出た目の和の期待値すなわち和の平均値を求めよ。ただし，このさいころのどの面の現れる見込みも同等であるとする。

（帯広畜産大）

（解） 目の和とその確率を表にすれば，次のようになる。

目の和	場合とその確率			
2	1と1	$\frac{1}{2}\times\frac{1}{2}$		
3	1と2	$\left(\frac{1}{2}\times\frac{1}{6}\right)\times2$		
4	1と3	$\left(\frac{1}{2}\times\frac{1}{6}\right)\times2$	2と2	$\frac{1}{6}\times\frac{1}{6}$
5	1と4	$\left(\frac{1}{2}\times\frac{1}{6}\right)\times2$	2と3	$\left(\frac{1}{6}\times\frac{1}{6}\right)\times2$
6	2と4	$\left(\frac{1}{6}\times\frac{1}{6}\right)\times2$	3と3	$\frac{1}{6}\times\frac{1}{6}$
7	3と4	$\left(\frac{1}{6}\times\frac{1}{6}\right)\times2$		
8	4と4	$\frac{1}{6}\times\frac{1}{6}$		

ゆえに，求める平均値は

$$2\times\frac{1}{4}+3\times\frac{1}{6}+4\times\frac{7}{36}+5\times\frac{2}{9}+6\times\frac{1}{12}$$
$$+7\times\frac{1}{18}+8\times\frac{1}{36}=4$$

答　4

＊　＊　＊

◆ **二項分布の平均**：

各試行において事象Eの起こる確率がp，起こらない確率を$q\ (=1-p)$としますと，この事象Eが起こる回数Xの確率分布は二項分布

$$P_r = {}_n\mathrm{C}_r p^r q^{n-r}$$

です。このとき，確率変数Xの平均値は

$$m = np$$

で与えられます。では，具体的な問題をやってみませんか。

練習 3. サイコロを10回振るとき，1の目が出る回数の平均値を求めよ。

ヒント 「サイコロを10回振る」というひとつの実験を何回もやるとき，1の目の出る回数の平均値はどうか，というのが問題の意味です。1の目の出る確率 $\dfrac{1}{6}$ $\left(出ない確率\ \dfrac{5}{6}\right)$ ですから，求める平均値は

$$10 \times \frac{1}{6} = \frac{5}{3}$$

です。

練習 4. 50歳の夫と48歳の妻が20年後まで生存する確率は，夫が0.15，妻が0.2である。現在，夫が50歳，妻が48歳である10組の夫婦が，20年後に夫婦のうち，少なくとも一方が生存している組の数をXとする。Xが二項分布にしたがう確率変数であるとして，Xの平均値を求めよ。

解 20年後に夫も妻も共に死亡している確率は $(1-0.15)(1-0.2)=0.85\times 0.8=0.68$ であるから，少なくも一方が生きている確率は 0.32 である。ゆえにXの平均値は

$$10 \times 0.32 = 3.2$$

である。　　　　答　3.2

＊　＊　＊

◆ 次に，平均値に関する練習を2，3やっておきましょう。まず，これです。

練習 5. 袋の中に20枚のカードが入っていて，それぞれ1，2，3，……，20なる数字が記入されている。最初に抽出した1枚の数値をX_1とし，残りから抽出した1枚の数値をX_2とする。X_1+X_2の平均値（期待値）を求めよ。

解 1，2，3，……，20から2つ選んで加えた数の期待値を求めればよい。その和は

$$\left.\begin{array}{r} 1+2,\ 1+3,\ 1+4,\ \cdots\cdots,\ 1+20 \\ 2+3,\ 2+4,\ \cdots\cdots,\ 2+20 \\ 3+4,\ \cdots\cdots,\ 3+20 \\ \cdots\cdots\cdots\cdots\cdots \\ 19+20 \end{array}\right\}\cdots(*)$$

で，その確率はいずれも $\dfrac{1}{{}_{20}\mathrm{C}_2}=\dfrac{1}{190}$ であるから，求める値は

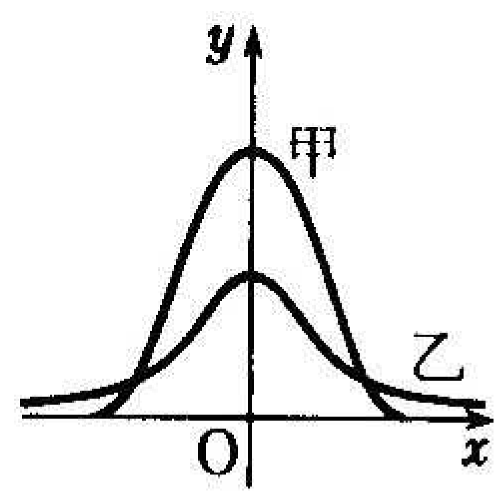
$$\frac{1}{190}\{(1+2+\cdots\cdots+20)\times 19\}=21$$

である。　　　　答　21

注 $(*)$の全部をバラバラにしてみると，1ないし20がいずれも19個ずつあることに注意。

練習 6. 右の図の甲，乙2つの度数分布曲線の平均値の大小を比較せよ。

（埼玉大）



答　等しい。

練習 7. n個の測定値 $x_1,\ x_2,\ \cdots\cdots,\ x_n$ の相加平均をm，標準偏差をσ，またcを定数とするとき，次のものを求めよ。

(1) $x_1+c,\ x_2+c,\ \cdots\cdots,\ x_n+c$ の相加平均

(2) $cx_1,\ cx_2,\ \cdots\cdots,\ cx_n$ の相加平均

(3) $cx_1{}^2,\ cx_2{}^2,\ \cdots\cdots,\ cx_n{}^2$ の相加平均

（高知大）

ヒント $m=\dfrac{1}{n}\sum x_i,\quad \sigma^2=\dfrac{1}{n}\sum(x_i-m)^2$

(1) $\dfrac{1}{n}\{\sum(x_i+c)\}=\dfrac{1}{n}\sum x_i+c=m+c$

(2) $\dfrac{1}{n}\sum cx_i=c\cdot\dfrac{1}{n}\sum x_i=cm$

(3) $\dfrac{1}{n}\sum cx_i{}^2=c\cdot\dfrac{1}{n}\sum\{(x_i-m)+m\}^2$

$$=\cdots\cdots=cm^2+c\sigma^2$$

(確率変数の)標準偏差とは何か

1回目　年　月　日
2回目　年　月　日
3回目　年　月　日

◆確率変数が与えられたときの標準偏差や分散を求めること，それが，ここの目的です。とくにめんどうはありませんよ。

◆ 確率変数Xのとり得る値x_1, x_2, …, x_nが $P(X=x_i)=p_i$ で与えられているとき，分散をσ^2やs^2や$V(X)$で表し，標準偏差はσやsや，s.d. や $D(X)$ で表します。

そして，

$$V(X)=\sum_{i=1}^{n}(x_i-m)^2p_i=E(X^2)-\{E(X)\}^2$$
$$D(X)=\sqrt{V(X)}$$

で与えられます。$E(X)$ はXの平均，$E(X^2)$ は X^2 の平均です。

*　*　*

◆ では，これをやってみませんか。

練習 1. Xは 1, 2, …, n を等しい確率でとる確率変数であって，分散は10である。このとき，nを求めよ。（東海大）

ヒント　$E(X)=\sum_{k=1}^{n}k\times\frac{1}{n}=\frac{1}{2}n(n+1)\times\frac{1}{n}$

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

$E(X^2)=\sum_{k=1}^{n}k^2\times\frac{1}{n}=\frac{1}{6}n(n+1)(2n+1)\times\frac{1}{n}$

$=\frac{(n+1)(2n+1)}{6}$

$\therefore\ V(X)=\frac{1}{6}(n+1)(2n+1)-\left\{\frac{1}{2}(n+1)\right\}^2$

$=10$

$\therefore\ n^2=121,\ n=11$　……答

練習 2. あるバレーボールの試合では，どちらかのチームが先に3セット勝てば試合は終了する。今，A，B両チームがこのバレーボールの試合を行うとき，1セットの勝負でAがBに勝つ確率をp，負ける確率をq（ただし，$p+q=1$とする）とし，pはこの試合を通して一定で，各セットの勝負は独立であるとする。このとき，次の問いに答えよ。

(1) この試合が4セットで終了する確率を求めよ。

(2) この試合が終了するまでに要するセット数をXとするとき，Xの確率分布を求めよ。

(3) この試合で $p=\frac{1}{2}$ とするとき，(2)のXの平均値と標準偏差を求めよ。（千葉大）

ヒント　(1) Aが勝って終了するのは，はじめの3回がAの2勝1負になるときですから，その確率は

$$({}_3C_2p^2\cdot q)\cdot p=3p^3q$$

同様にして，Bが勝って終了する確率も$3pq^3$ですから，4セットで終了する確率は

$$3p^3q+3pq^3=3pq(p^2+q^2)\ \text{……答}$$

となります。

(2) 3勝すれば試合が終わるので

$$X=3,\ 4,\ 5$$

で，その確率分布は下のようになります。

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

(3) $p=q=\frac{1}{2}$ ですから，(2)の確率分布表は，次のようになる。

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

したがって

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

$$E(X^2)=3^2\times\frac{2}{8}+4^2\times\frac{3}{8}+5^2\times\frac{3}{8}=\frac{141}{8}$$

$$\therefore\quad V(X)=E(X^2)-\{E(X)\}^2$$
$$=\frac{141}{8}-\left(\frac{33}{8}\right)^2=\frac{39}{64}$$

$$\therefore\quad D(X)=\sqrt{V(X)}=\frac{\sqrt{39}}{8}\cdots\cdots\text{答}$$

練習 3. 数字0，1，2を1つずつ記入した等形，等大，等質の3枚のカードの中から任意に1枚を抜きとってカードに記入された数字Xを読む。次にそのカードをもとに戻して，さらに1枚を抜きとって，カードに記入された数字Yを読む。このとき，Z=XY として，

(1) Zのとるそれぞれの値に対する確率を求めよ。

(2) それぞれの確率変数Zの平均値と分散を求めよ。 (大分大)

ヒント (1) ZとX，Yの値を表にしてみますと次のようになります。

Z	0					1	2		4
X	0	0	1	2	0	1	1	2	2
Y	0	1	0	0	2	1	2	1	2

したがって，Zの確率分布表は下のようになります。

Z	0	1	2	4	計
$P(Z)$	$\frac{5}{9}$	$\frac{1}{9}$	$\frac{2}{9}$	$\frac{1}{9}$	1

(2) $E(Z)=0\times\frac{5}{9}+1\times\frac{1}{9}+2\times\frac{2}{9}+4\times\frac{1}{9}$

$=1 \qquad\cdots\cdots$答

$E(Z^2)=0^2\times\frac{5}{9}+1^2\times\frac{1}{9}+2^2\times\frac{2}{9}+4^2\times\frac{1}{9}$

$=\frac{25}{9}$

$\therefore\quad V(Z)=E(Z^2)-\{E(Z)\}^2=\frac{25}{9}-1=\frac{16}{9}$

答 $\frac{16}{9}$

* * *

◆ では，やや総合的な問題を練習してみませんか。

練習 4. A，B 2つの箱があり，Aには赤球が2個，Bには白球が2個入っている。いま，両方の箱から同時に1個ずつとり出して，とり出した球を入れかえることにする。この試行をn回くり返したときのAの中にある赤球の個数を確率変数 X_n とし，X_n が 0，1，2 の値をとる確率をそれぞれ p_n，q_n，r_n とする。

(1) p_n，q_n，r_n の値を求めよ。

(2) X_n の分散 $V(X_n)$ を求めよ。

(静岡大)

ヒント (1) n回目にAの赤球が0個になるのは $(n-1)$ 回目にA，Bとも赤，白の球が各々1個の状態になっていて，次の操作で，Aの赤とBの白が交換される場合ですから

$$p_n=q_{n-1}\times\frac{1}{2}\times\frac{1}{2}=\frac{1}{4}q_{n-1}\cdots\cdots①$$

同じようにして

$$r_n=\frac{1}{4}q_{n-1}\qquad\cdots\cdots②$$

また，$q_n=1-(p_n+r_n)$ より

$$q_n=1-\frac{1}{2}q_{n-1}\qquad\cdots\cdots③$$

③より

$$q_n=\frac{1}{3}\left\{2+\left(-\frac{1}{2}\right)^{n-1}\right\}$$

したがって

$$p_n=r_n=\frac{1}{12}\left\{2+\left(-\frac{1}{2}\right)^{n-2}\right\}$$

(2) $E(X_n)=0\cdot p_n+1\cdot q_n+2\cdot r_n=1$

$$\therefore\quad V(X_n)=(0-1)^2p_n+(1-1)^2q_n+(2-1)^2r_n$$
$$=p_n+r_n$$
$$=\frac{1}{6}\left\{2+\left(-\frac{1}{2}\right)^{n-2}\right\}$$

となりましょう。本問に関係はありませんが，これから $\lim_{n\to\infty}V(X_n)=\frac{1}{3}$ となります。

それにしても無限回くり返すとはなにか。

確率変数の変換

1回目　年　月　日
2回目　年　月　日
3回目　年　月　日

◆資料の平均や分散を求めるのに仮平均を使うと便利であった。確率変数の場合にも同じことが起こるのです。

◆ ここでは確率変数を変換すると，平均値や標準偏差がどのように変わるかを考えるのが目的です。

練習 1. 確率変数Xの平均が $E(X)=-3$ で分散が $V(X)=5$ であるとする。確率変数 $Y=2X+5$ とするとき，Yの平均は $E(Y)=\boxed{}$ で，分散は $V(Y)=\boxed{}$ である。　(東海大)

ヒント　確率変数Xのとり得る値が $x_1, x_2, \cdots\cdots, x_n$ で $P(X=x_i)=p_i$ $(i=1, 2, \cdots\cdots, n)$ としますと，Yのとり得る値は

$y_1=2x_1+5,\ y_2=2x_2+5,\ \cdots,\ y_n=2x_n+5$

ですから

$$E(Y)=\sum_{i=1}^{n}p_iy_i=\sum_{i=1}^{n}p_i(2x_i+5)$$
$$=2\sum_{i=1}^{n}p_ix_i+5\sum_{i=1}^{n}p_i$$
$$=2E(X)+5=2\cdot(-3)+5=-1$$

となります。また，

$$V(Y)=\sum_{i=1}^{n}p_i(y_i-(-1))^2$$
$$=\sum_{i=1}^{n}p_i(2x_i+6)^2=\sum_{i=1}^{n}p_i\{2(x_i+3)\}^2$$
$$=4\sum_{i=1}^{n}p_i(x_i-(-3))^2=4V(X)$$
$$=4\cdot5=20$$

答　-1, 20

注　ここでは，上のように，平均値および分散の定義から直接計算しましたが，実は次のような公式があります。

$$E(aX+b)=aE(X)+b$$
$$\{\sigma(aX+b)\}^2=a^2\{\sigma(X)\}^2$$

となります。EとVのちがいに，くれぐれも御用心。

この公式を使うなら，上の練習1では

$$E(2X+5)=2E(X)+5=2(-3)+5=-1$$
$$V(2X+5)=2^2V(X)=4\cdot5=20$$

となって，すぐ求められます。

次も同じです。

練習 2. 確率変数Xの平均がm，標準偏差がσ $(\sigma>0)$ のとき，確率変数 $Y=aX+b$ の平均が0，標準偏差が1となるように，定数a，bの値を求めよ。ただし，$a>0$ とする。

ヒント　$E(Y)=aE(X)+b=am+b$

$$\sigma(Y)=a\sigma(X)=a\sigma$$

ですから

$$am+b=0\ かつ\ a\sigma=1$$

$$\therefore\quad a=\frac{1}{\sigma},\ b=-\frac{m}{\sigma}$$　……答

練習 3. 確率変数Xの平均が3，標準偏差が $\sqrt{2}$ のとき，確率変数 $Y=aX+b$ $(a>0)$ の平均が0，分散が1となるように定数a，bを求めよ。　(宮崎大)

ヒント　上とまったく同じです。

$$E(Y)=aE(X)+b=3a+b=0$$
$$\sigma(Y)=a\sigma(X)=a\cdot\sqrt{2}=1$$

より

$$a=\frac{1}{\sqrt{2}},\ b=-\frac{3}{\sqrt{2}}$$　……答

*　　*　　*

◆ ではやや，総合的な問題を練習してみませんか。

練習 4. 1枚の硬貨を4回投げるとき，表の出る回数から裏の出る回数を引いた数を確率変数Xで表し，$Y=\dfrac{X}{2}+3$ とおく。このとき

(1)　Yの確率分布，平均（期待値）および分散を求めよ。

(2)　無限級数 $\displaystyle\sum_{n=1}^{\infty}\left\{\frac{Y(a-Y^2)}{40}\right\}^n$ が収束する確率が $\dfrac{1}{8}$ となるように，実数aのとるべき値の範囲を定めよ。　(岡山大)

ヒント　(1)　硬貨を4回投げて出るXやYの確率分布を調べると次の表のようです。

表	0	1	2	3	4	計
裏	4	3	2	1	0	
X	-4	-2	0	2	4	
Y	1	2	3	4	5	
P	$\frac{1}{16}$	$\frac{4}{16}$	$\frac{6}{16}$	$\frac{4}{16}$	$\frac{1}{16}$	1

そこでYの平均値は

$$\frac{1}{16}(1\times1+2\times4+3\times6+4\times4+5\times1)$$
$$=3$$

です。また，分散は

$$\frac{1}{16}((-2)^2\times1+(-1)^2\times4+0^2\times6+1^2\times4+2^2\times1)=1$$

です。

(2)　$\displaystyle\sum_{n=1}^{\infty}\left\{\frac{Y(a-Y^2)}{40}\right\}^n$ が収束する確率が $\dfrac{1}{8}$ となるためには

$Y=1$，5 について

$$\frac{|Y(a-Y^2)|}{40}<1$$

が成り立ち，それ以外について

$$\frac{|Y(a-Y^2)|}{40}\geqq1$$

が成り立てばよい。すなわち

$|1\cdot(a-1)|<40$，$|2\cdot(a-4)|\geqq40$，

$|3\cdot(a-9)|\geqq40$，$|4\cdot(a-16)|\geqq40$，

$|5\cdot(a-25)|<40$

を満足するaを求めればよいでしょう。これを解くと

$$26\leqq a<33\qquad\cdots\cdots答$$

*　*　*

練習5.　Xを確率変数とし，$X=x_k$ となる確率を p_k とする。$Y=ax_k+b$（a，b は定数で，$a\neq0$）となる確率が p_k であるとき，確率変数を $Y=aX+b$ と書く。ただし，$k=1,\ 2,\ \cdots\cdots,N$ とし，$\displaystyle\sum_{k=1}^{N}p_k=1$ とする。

(1)　正しい硬貨を5回投げて，表の出た回数と裏の出た回数の差の絶対値を X_0 とする。確率変数 X_0 の平均値と分散を求めよ。

(2)　(1)の X_0 から出発して，$3X_n=X_{n-1}+4$（$n=1,\ 2,\ \cdots\cdots$）によって定義される確率変数の列 $X_0, X_1, X_2, \cdots\cdots, X_n, \cdots\cdots$ がある。X_n の平均値と分散を求めよ。

(高知医大)

ヒント　(1)　5回中n回表が出る確率は

$$ {}_5\mathrm{C}_n\left(\frac{1}{2}\right)^5$$

ですから

$$E(X_0)=1\cdot\frac{5}{2^3}+3\cdot\frac{5}{2^4}+5\cdot\frac{1}{2^4}$$
$$=\frac{15}{8}\qquad\cdots\cdots答$$

$$\{\sigma(X_0)\}^2=\left(1-\frac{15}{8}\right)^2\cdot\frac{5}{2^3}+\left(3-\frac{15}{8}\right)^2\cdot\frac{5}{2^4}+\left(5-\frac{15}{8}\right)^2\cdot\frac{1}{2^4}=\frac{95}{64}\qquad\cdots\cdots答$$

(2)　$3E(X_n)=E(X_{n-1})+4$

ですから，これを変形して

$$3(E(X_n)-2)=E(X_{n-1})-2$$
$$\therefore\quad E(X_n)-2=\{E(X_0)-2\}\left(\frac{1}{3}\right)^n$$
$$=-\frac{1}{8}\left(\frac{1}{3}\right)^n$$
$$\therefore\quad E(X_n)=2-\frac{1}{8}\left(\frac{1}{3}\right)^n\qquad\cdots\cdots答$$

分散は $\{\sigma(X_n)\}^2=\dfrac{1}{9}\{\sigma(X_{n-1})\}^2$

ところが $\{\sigma(X_0)\}^2=\dfrac{95}{64}$ ですから

$$\{\sigma(X_n)\}^2=\frac{95}{64}\left(\frac{1}{9}\right)^n\qquad\cdots\cdots答$$

○確率変数の和と積の統計量

1回目　年　月　日
2回目　年　月　日
3回目　年　月　日

◆独立ないくつかの確率変数の和や積の平均や分散をもとの統計量から求められるものもあります。

◆ 試行Aについての確率変数をX，試行Bについての確率変数をYとして，その確率分布が下のようであったとしましょう。

X	x_1	x_2	計
P	p_1	p_2	1

Y	y_1	y_2	計
P	q_1	q_2	1

そうしますと変数X，Yの積とは

$$\{x_1y_1,\ x_1y_2,\ x_2y_1,\ x_2y_2\}$$

からなり，そのおのおのの確率はそれぞれ

$$\{p_1q_1,\ p_1q_2,\ p_2q_1,\ p_2q_2\} \quad \cdots(*)$$

となります。したがってその平均は

$$\begin{aligned}&(x_1y_1)(p_1q_1)+(x_1y_2)(p_1q_2)\\&+(x_2y_1)(p_2q_1)+(x_2y_2)(p_2q_2)\\=&(x_1p_1+x_2p_2)(y_1q_1+y_2q_2)\\=&E(X)\cdot E(Y)\end{aligned}$$

となります。ただし，(*) のところで，**XとYが独立である**ことが大切です。

これを式で表して

$$\boldsymbol{E(XY)=E(X)E(Y)}$$

としましょう。では：――さっそくながら次をやってみませんか。

■練習 1. 2個のさいころを同時に投げるとき，出る目の積の平均を求めよ。

ヒント　2つのさいころの出る目をX，Yとしますと積の値は右の表のようになります。そしておのおのの生起する確率はいずれも$\frac{1}{36}$

X＼Y	1	2	3	4	5	6
1	1	2	3	4	5	6
2	2	4	6	8	10	12
3	3	6	9	12	15	18
4	4	8	12	16	20	24
5	5	10	15	20	25	30
6	6	12	18	24	30	36

ですから，これらの数値の和に$\frac{1}{36}$を掛けたものです。

しかし，左にあげた公式を使うなら

$$\begin{aligned}E(XY)&=E(X)E(Y)\\&=\left\{\frac{1}{6}(1+2+3+4+5+6)\right\}^2\\&=\left(\frac{7}{2}\right)^2=\frac{49}{4}\end{aligned}$$

……答

■練習 2. 60人の生徒について，数学の成績をx，英語の成績をyとするとき，(x, y)の組は次の4組だけに分類されて

(80, 90)，(60, 50) がともにa人

(60, 80)，(80, 60) がともにb人

であるとする。

この60人の生徒のうちから無作為に1人を抽出するとき，数学の成績をX，英語の成績をYとする。

(1) $E(X)\cdot E(Y)=E(XY)$ となるようなa，bを求めよ。ただし，$E(Z)$ は確率変数Zの平均（期待値）を表す。

(2) 上の(1)の場合X，Yは独立であるかどうか調べよ。ただし，確率変数X，Yが独立であるとは

Xのとる値が $x_1,\ x_2,\ \cdots\cdots,\ x_n$

Yのとる値が $y_1,\ y_2,\ \cdots\cdots,\ y_m$

であるとき，任意のi，j $(i=1,\ 2, \cdots,\ n\ ;\ j=1,\ 2,\ \cdots,\ m)$ に対して

事象 $X=x_i$ と事象 $Y=y_i$ とが独立である

ことである。

ヒント　(1)

$$E(X)=\frac{1}{2(a+b)}(80a+60a+60b+80b)=70$$

$$E(Y)=\frac{1}{2(a+b)}(90a+50a+80b+60b)=70$$

次に XY の確率分布は下のようになります。

XY	3000	4800	7200	和
P	$\frac{a}{60}$	$\frac{2b}{60}$	$\frac{a}{60}$	1

$$\therefore\quad E(XY)=\frac{100}{60}(30a+2b\cdot48+72a)$$

$$=10(17a+16b)$$

$$\therefore\quad 10(17a+16b)=70\times70$$

$$17a+16b=490 \quad\cdots\cdots①$$

また $$a+b=30 \quad\cdots\cdots②$$

$$\therefore\quad a=10,\ b=20$$

(2) $$P_{X=80}(Y=90)=\frac{a}{a+b}=\frac{10}{30}=\frac{1}{3}$$

$$P(Y=90)=\frac{a}{60}=\frac{1}{6}$$

$$\therefore\quad P_{X=80}(Y=90)\fallingdotseq P(Y=90)$$

よって，X，Yは独立でないことがわかります。

* * *

◆ 確率変数X，Yが独立のとき

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

という関係が成り立ちます。

その証明は

$$\{\sigma(X+Y)\}^2=E((X+Y)^2)-\{E(X+Y)\}^2$$
$$=E(X^2+2XY+Y^2)-\{E(X)+E(Y)\}^2$$
$$=E(X^2)+2E(XY)+E(Y^2)$$
$$-[\{E(X)\}^2+2E(X)E(Y)+\{E(Y)\}^2]$$

ところがX，Yが独立ですから

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

$$\therefore\quad \{\sigma(X+Y)\}^2=[E(X^2)-\{E(X)\}^2]$$
$$+[E(Y^2)-\{E(Y)\}^2]$$
$$=\{\sigma(X)\}^2+\{\sigma(Y)\}^2$$

では，具体的な練習へいきましょう。

■練習 3. 2個のさいころを同時に投げるとき，出る目の和の分散を求めよ。

ヒント さいころを投げて出る目をXとすると き，Xの確率分布から分散を求めると次のようになります。(定義にしたがってやると)

X	1	2	3	4	5	6	計
p	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	1
pX	$\frac{1}{6}$	$\frac{2}{6}$	$\frac{3}{6}$	$\frac{4}{6}$	$\frac{5}{6}$	$\frac{6}{6}$	$\frac{7}{2}$
$X-\frac{7}{2}$	$-\frac{5}{2}$	$-\frac{3}{2}$	$-\frac{1}{2}$	$\frac{1}{2}$	$\frac{3}{2}$	$\frac{5}{2}$	
$\left(X-\frac{7}{2}\right)^2$	$\frac{25}{4}$	$\frac{9}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{9}{4}$	$\frac{25}{4}$	
$p\left(X-\frac{7}{2}\right)^2$	$\frac{25}{24}$	$\frac{9}{24}$	$\frac{1}{24}$	$\frac{1}{24}$	$\frac{9}{24}$	$\frac{25}{24}$	$\frac{35}{12}$

つまり，1つのさいころを投げるとき，出る目の分散は $\frac{35}{12}$ です。そこで，2つのさいころを投げて出る目の数をそれぞれX，Yとしますと，XとYは独立ですから，目の和の分散は

$$\{\sigma(X+Y)\}^2=\{\sigma(X)\}^2+\{\sigma(Y)\}^2$$
$$=\frac{35}{12}+\frac{35}{12}=\frac{35}{6}$$

です。

■練習 4. 100円硬貨1枚と10円硬貨2枚を投げて表の出た硬貨の金額の合計をsとするとき，sの標準偏差を求めよ。

ヒント 100円硬貨および2枚の10円硬貨を投げて表が出たときに得られる金額をそれぞれX，Y，Zとすると

$$\sigma^2(X)=\frac{1}{2}\times100^2-50^2=2500$$

$$\sigma^2(Y)=\sigma^2(Z)=\frac{1}{2}\times10^2-5^2=25$$

X，Y，Zは独立であるから

$$\sigma^2(X+Y+Z)=\sigma^2(X)+\sigma^2(Y)+\sigma^2(Z)$$
$$=2500+25+25$$
$$=2550$$

$$\therefore\quad \sigma(X+Y+Z)=\sqrt{2550}$$
$$=5\sqrt{102} \quad\cdots\cdots 答$$

なお，$5\sqrt{102}\fallingdotseq50.50$ です。

◐二項分布とは何か

1回目　年　月　日
2回目　年　月　日
3回目　年　月　日

◆確率の好きな人でも統計はイヤな人が多い。そして，統計の中でもっともよく現れるのは正規分布と二項分布なんですよ。

◆ 1回の試行で事象Aが起こる確率を p としますと，

n 回の試行で事象Aの起こる回数 X を確率変数にとると，その変数 X が r となる値をとる確率は

$${}_nC_r p^r q^{n-r} \quad (q=1-p)$$

となります。この確率分布を **二項分布** といいます。では，さっそくこれを：——

■練習 1. サイコロを5回振って偶数の目の出る回数を確率変数にとるとき，その確率分布を求めよ。

(ヒント) つまりサイコロを5回振り，偶数の目が0回，1回，……，5回出る確率を求めよ，というわけだ。$X=r$ の確率は ${}_5C_r\left(\frac{1}{2}\right)^5$ だから，

X	0	1	2	3	4	5
p_r	$\frac{1}{32}$	$\frac{5}{32}$	$\frac{10}{32}$	$\frac{10}{32}$	$\frac{5}{32}$	$\frac{1}{32}$

となります。分母を32にしたほうが見やすいので約してありませんが，もちろん約してもよいのです。

■練習 2. サイコロを10回振るとき，1の目が何回出る場合がもっとも確からしいか。

(ヒント) 10回のうち r 回1の目が出る確率を W_r としますと

$$W_r={}_{10}C_r\left(\frac{1}{6}\right)^r\left(\frac{5}{6}\right)^{10-r}={}_{10}C_r\frac{5^{10-r}}{6^{10}}$$

ゆえに W_r を最大にする r を求めるには

$${}_{10}C_r5^{10-r}$$

を最大にする r を求めればよいわけ。さて，

$$\frac{W_{r+1}}{W_r}=\frac{{}_{10}C_{r+1}5^{10-(r+1)}}{{}_{10}C_r5^{10-r}}=\frac{{}_{10}C_{r+1}5^r}{{}_{10}C_r5^{r+1}}$$

$$=\frac{10\cdot9\cdots\cdots(11-r)(10-r)}{1\cdot2\cdots\cdots r(r+1)}\times\frac{1\cdot2\cdots\cdots r}{10\cdot9\cdots(11-r)}\times\frac{1}{5}$$

$$=\frac{10-r}{r+1}\times\frac{1}{5}$$

$$\therefore\quad (10-r)-5(r+1)=5-6r<0$$

$$\therefore\quad W_1>W_2>\cdots\cdots$$

で，1回起こることがもっとも確からしいということがわかります。

[答] 1回

* * *

◆ 二項分布で大切なことのひとつは **平均** と **標準偏差** を求めることです。

■練習 3. 1回の試行で事象Aの起こる確率を p，起こらない確率を $q(=1-p)$ とする。n 回の試行で，この事象Aの起こる回数を x とするとき，x の平均を求めよ。

(解) $x=r$ となる確率は ${}_nC_rp^rq^{n-r}$ であるから，x の平均 $E(x)$ は

$$E(x)=1{}_nC_1pq^{n-1}+2{}_nC_2p^2q^{n-2}+\cdots\cdots$$
$$\cdots+n{}_nC_np^n$$

$$=1\cdot\frac{n}{1}pq^{n-1}+2\cdot\frac{n(n-1)}{1\cdot2}p^2q^{n-2}+\cdots$$
$$\cdots+n\cdot\frac{n(n-1)\cdots\cdots1}{1\cdot2\cdots\cdots n}p^n$$

$$=npq^{n-1}+\frac{n(n-1)}{1}p^2q^{n-2}+\cdots$$
$$+\frac{n(n-1)\cdots\cdots1}{1\cdots\cdots(n-1)}p^n$$

$$=np\left\{q^{n-1}+\frac{n-1}{1}pq^{n-2}+\cdots\right.$$
$$\left.\cdots+\frac{(n-1)\cdots\cdots1}{1\cdots\cdots(n-1)}p^{n-1}\right\}$$

$$=np(p+q)^{n-1}=np\cdot1^{n-1}=np$$

[答] np

(注) わからない人は **二項定理**（☞ p.44）を復習してください。

練習 4. 1回の試行で事象Aの起こる確率をp，起こらない確率を $q(=1-p)$ とする。n回の試行において，この事象Aの起こる回数をxとするとき，xの標準偏差を求めよ。

(ヒント) 標準偏差σというのは x_1, x_2, ……, x_n の平均を$\overline{x}$とするとき

$$\sigma=\sqrt{\frac{(x_1-\overline{x})^2+(x_2-\overline{x})^2+\cdots+(x_n-\overline{x})^2}{n}}$$

のことでした。根号内を変形すると

$$\sqrt{\frac{x_1{}^2+x_2{}^2+\cdots+x_n{}^2}{n}-\left(\frac{x_1+x_2+\cdots+x_n}{n}\right)^2}$$

とも書けます。ところが二項分布の場合には

$\dfrac{x_1+x_2+\cdots+x_n}{n}=np$ なんですから

$$\sqrt{\frac{x_1{}^2+x_2{}^2+\cdots+x_n{}^2}{n}-(np)^2}$$

と書けましょう。結局，根号内第1項の計算が焦点になってきたわけです。では：——

(解) xの2乗平均を $E(x^2)$ で表すと

$$E(x^2)=1^2{}_n\mathrm{C}_1pq^{n-1}+2^2{}_n\mathrm{C}_2p^2q^{n-2}+\cdots$$
$$\cdots+n^2{}_n\mathrm{C}_np^n \quad \cdots\cdots①$$

ところが

$$r{}_n\mathrm{C}_r=r\cdot\frac{n(n-1)\cdots(n-r+2)(n-r+1)}{1\cdot2\cdot3\cdots(r-1)\cdot r}$$
$$=n\cdot\frac{(n-1)(n-2)\cdots(n-r+1)}{1\cdot2\cdot\cdots\cdot(r-1)}$$
$$=n{}_{n-1}\mathrm{C}_{r-1}$$

であるから（ヤレヤレ，タイテイココデ混乱!!）

$$(r-1){}_{n-1}\mathrm{C}_{r-1}=(n-1){}_{n-2}\mathrm{C}_{r-2}$$

でもある。（初心者ハココデヤメテモイイゼ）

$$\therefore\quad r^2{}_n\mathrm{C}_r=r(r{}_n\mathrm{C}_r)=r(n{}_{n-1}\mathrm{C}_{r-1})$$
$$=n(r{}_{n-1}\mathrm{C}_{r-1})=n\{(r-1)+1\}{}_{n-1}\mathrm{C}_{r-1}$$
$$=n(r-1){}_{n-1}\mathrm{C}_{r-1}+n{}_{n-1}\mathrm{C}_{r-1}$$
$$=n(n-1){}_{n-2}\mathrm{C}_{r-2}+n{}_{n-1}\mathrm{C}_{r-1}$$
$$\therefore\quad r^2{}_n\mathrm{C}_rp^rq^{n-r}$$
$$=n(n-1)p^2{}_{n-2}\mathrm{C}_{r-2}p^{r-2}q^{n-r}$$
$$+np{}_{n-1}\mathrm{C}_{r-1}p^{r-1}q^{n-r}$$

これを①に代入して

$$E(x^2)=n(n-1)p^2\{{}_{n-2}\mathrm{C}_0q^{n-2}$$
$$+{}_{n-2}\mathrm{C}_1pq^{n-3}+\cdots\cdots+{}_{n-2}\mathrm{C}_{n-2}p^{n-2}\}$$
$$+np\{{}_{n-1}\mathrm{C}_0q^{n-1}+{}_{n-1}\mathrm{C}_1pq^{n-2}+\cdots$$
$$\cdots+{}_{n-1}\mathrm{C}_{n-1}p^{n-1}\}$$
$$=n(n-1)p^2(p+q)^{n-2}+np(p+q)^{n-1}$$
$$=n(n-1)p^2+np=np(np-p+1)$$

ゆえに，求める標準偏差は

$$\sigma^2=E(x^2)-E(x)^2$$
$$=np(np-p+1)-(np)^2=npq$$
$$\therefore\quad \sigma=\sqrt{npq} \qquad \cdots\cdots\boxed{答}$$

(注) とてもこの計算は覚えられそうにもありませんね。結果だけ覚えておいて使えるようにしておくことが肝心です。

*　*　*

◆ では，具体的な問題をやってみませんか。

練習 5. 飛行機のエンジンが飛行中故障を起こす確率はqで，エンジンの少なくも半数が動いていれば，飛行を続けることができるものとする。また，1つのエンジンの故障は他のエンジンに影響を与えないものとする。2つのエンジンの飛行機のほうが，4つのエンジンの飛行機より安全なのは，qがどんな値のときか。（名古屋市大）

(ヒント) 1つのエンジンが故障を起こさない確率は $p=1-q$ です。

さて，2つのエンジンをもつ飛行機が飛行可能である確率は，故障を起こさないエンジンが2個のときと1個のときで，その確率は

$1-$（2個とも故障を起こす確率）

$=1-q^2$

です。次に4つのエンジンをもつときには飛行可能の確率は

$1-$(故障を起こさないもの0の確率)

$-$(故障を起こさないもの1個の確率)

$=1-q^4-4pq^3=1-4q^3+3q^4$

そこで $1-q^2>1-4q^3+3q^4$ を解いて求める答は $q>\dfrac{1}{3}$ とわかります。

① 二項分布の平均値

1回目　年　月　日
2回目　年　月　日
3回目　年　月　日

◆二項分布 ${}_nC_rp^rq^{n-r}$ の平均値はnpですが，その扱い方が中心です。そして，それに関連することも，です。

◆　1回の試行で事象Eの起こる確率をpとしましょう。この試行をn回独立にくり返すとき，事象Eの起こる回数を確率変数Xとしますと，$X=x$ となる確率 $P(X=x)$ は

$$P(X=x)={}_nC_xp^xq^{n-x}$$

$$(x=0,\ 1,\ \cdots\cdots,\ n\ ;\ q=1-p)$$

で与えられます。この確率分布を**二項分布**といい，$\mathrm{B}(n,\ p)$ で表します。

*　*　*

■**練習 1.** では，次の確率分布を$\mathrm{B}(n,\ p)$の形式で表しましょう。

1つのさいころを，次の回数だけ振ったとき，6の目の出る回数Xを確率変数とする。

(1)　30回　(2)　300回　(3)　3000回

(ヒント)　さいころを振ったとき，6の目の出る確率pは

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

ですから，

$$q=\frac{5}{6}$$

したがって(1)，(2)，(3)はそれぞれ

$$\mathrm{B}\left(30,\ \frac{1}{6}\right),\ \mathrm{B}\left(300,\ \frac{1}{6}\right),\ \mathrm{B}\left(3000,\ \frac{1}{6}\right)$$

で表されます。

*　*　*

◆　ところで$\mathrm{B}(n,\ p)$の平均値（期待値）は次のようになります。

$$E(X)=np \qquad \cdots\cdots(*)$$

この証明は140ページにありますが，ここでは，別解をあげておきましょう。

(*)の証明：二項分布$\mathrm{B}(n,\ p)$において，$X=r$ となる確率は ${}_nC_rp^rq^{n-r}$ ですから，Xの平均 $E(X)$ は

$$E(X)=1{}_nC_1pq^{n-1}+2{}_nC_2p^2q^{n-2}+\cdots\cdots$$
$$\cdots\cdots+n{}_nC_np^n$$

で与えられます。

ところが，二項展開

$$(q+px)^n=q^n+{}_nC_1q^{n-1}px+{}_nC_2q^{n-2}(px)^2$$
$$+\cdots\cdots+{}_nC_n(px)^n$$

は，xについて恒等式ですから，xについて微分しますと

$$np(q+px)^{n-1}=1{}_nC_1q^{n-1}p+2{}_nC_2q^{n-2}p^2x$$
$$+\cdots\cdots+n{}_nC_np^nx^{n-1}$$

となります。ここで $x=1$ とおくと

$$np(q+p)^{n-1}=1\cdot{}_nC_1pq^{n-1}+2\cdot{}_nC_2p^2q^{n-2}$$
$$+\cdots\cdots+n{}_nC_np^n$$

ところが $p+q=1$ ですから

$$1\cdot{}_nC_1pq^{n-1}+2\cdot{}_nC_2p^2q^{n-2}+\cdots\cdots+n{}_nC_np^n$$
$$=np$$

$$\therefore\quad E(X)=np$$

となります。

そこで，次を：——

■**練習 2.** $\mathrm{B}\left(10,\ \frac{1}{5}\right)$ について平均を求めよ。

(解)　$n=10,\ p=\frac{1}{5}$ ですから

$$E(X)=10\cdot\frac{1}{5}=2 \qquad \cdots\cdots\text{答}$$

■**練習 3.** 硬貨を20回投げるとき，表が出る回数の平均を求めよ。

(ヒント)　表が出る確率も裏が出る確率も $\frac{1}{2}$ ですから，X回出る確率は ${}_{20}C_X\left(\frac{1}{2}\right)^X\left(\frac{1}{2}\right)^{20-X}$ で与えられます。つまり，≪二項分布

$$\mathrm{B}\left(20,\ \frac{1}{2}\right)$$

に従う≫のです。

$$\therefore\quad E(X)=np=20\cdot\frac{1}{2}=10$$

■練習 4. 白球3個，黒球7個が入っている袋から1個をとり出してもとに戻すことを4回くり返すとき，白球が出る回数をXとする。Xの平均値を求めよ。

(ヒント) 白球が出る回数 $X=x$ としますと

$$p_x={}_4\mathrm{C}_x\left(\frac{3}{10}\right)^x\left(\frac{7}{10}\right)^{4-x}\quad(x=0,1,2,3,4)$$

となりますから，確率変数Xは ≪二 項 分 布 $\mathrm{B}\left(4,\ \frac{3}{10}\right)$ に従う≫のです。

ゆえに，平均値

$$E(X)=4\cdot\frac{3}{10}=\frac{6}{5}$$

となります。

■練習 5. 4枚の硬貨を同時に投げる試行を4回くり返すとき，2枚が表で2枚が裏となる回数Xの確率分布の式を求めよ。また，Xの平均値を求めよ。　(弘前大)

(ヒント) 4枚の硬貨を同時に投げて2枚が表で2枚が裏である確率は

$$_4\mathrm{C}_2\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^2$$

$$=\frac{4\cdot3}{1\cdot2}\cdot\frac{1}{16}=\frac{3}{8}$$

で与えられます。したがって

$$P(X=x)={}_4\mathrm{C}_x\left(\frac{3}{8}\right)^x\left(\frac{5}{8}\right)^{4-x}$$

$$(x=0,\ 1,\ 2,\ 3,\ 4)$$

つまり≪二項分布

$$\mathrm{B}\left(4,\ \frac{3}{8}\right)$$

に従う≫のです。そして平均値は

$$4\cdot\frac{3}{8}=\frac{3}{2}$$

■練習 6. ある機械で生産される製品のうち，10%は不良品であるという。この機械で10個の製品を作るとき，そのうちの不良品の数を S_{10} とする。このとき，次のものを求めよ。

(1) $P(S_{10}=8)$

(2) $E(S_{10})$

(ヒント) S_{10} の分布は二項分布 B(10, 0.1) に従います。したがって

$$P(S_{10}=x)={}_{10}\mathrm{C}_x(0.1)^x(0.9)^{10-x}$$

そこで

(1) $P(S_{10}=8)={}_{10}\mathrm{C}_8(0.1)^8(0.9)^2$

$$=\frac{10\cdot9}{1\cdot2}\times\frac{9^2}{10^{10}}=\frac{3645}{10^{10}}$$

(2) $E(S_{10})=np=10\times0.1=1$

*　*　*

◆ では，やや総合的な問題をやってみませんか。

■練習 7. 大小2個のさいころを同時に投げる試行をTとする。1回の試行Tでさいころが2個とも偶数の目の出る事象をAとする。試行Tをくり返すとき，n回目に事象Aが起これば $X_n=1$，事象Aが起こらなければ $X_n=0$ とし，$S=X_1+X_2+X_3+X_4$ とする。このとき，Sの期待値（平均値）$E(S)$ を求めよ。　(北大)

(ヒント) 大小2個のさいころを同時に投げて2個とも偶数の目の出る確率は

$$\frac{1}{2}\times\frac{1}{2}=\frac{1}{4}$$

です。したがって，確率変数Sは二項分布 $\mathrm{B}\left(4,\ \frac{1}{4}\right)$ に従います。ゆえにその期待値は

$$np=4\times\frac{1}{4}=1$$

となります。

(注) ちょっとみるとメンドウにみえますがわかってしまえば，コトモナシというわけ。

*　*　*

◆ このように二項分布の平均値の扱いにめんどうはありません，むしろ問題の読み方にポイントのあることが多いのです。

○二項分布と標準偏差

1回目　年　月　日
2回目　年　月　日
3回目　年　月　日

◆二項分布は，あんがい，身近なところに多いものです。だから，いやがらず，おそれず，これになじむことが大切でしょう。

◆ 二項分布 B$(n,\ p)$ の確率分布表は次のようです。

x	0	1	2	…	n	計
$P(X=x)$	${}_n\mathrm{C}_0q^n$	${}_n\mathrm{C}_1pq^{n-1}$	${}_n\mathrm{C}_2p^2q^{n-2}$	…	${}_n\mathrm{C}_np^n$	1

ところで，B$(n,\ p)$ の平均値は np で与えられますが，分散は npq，標準偏差は $\sqrt{npq}$ で与えられます。この証明は141ページにあげてありますが，ここでは別解をあげておきましょう。だが，その前に，これをやっておくとしましょう。

練習1. 二項分布 B$(2,\ p)$ について，分散を定義にしたがってを求めよ。

(解) 確率分布表は

X	0	1	2	計
p_r	q^2	$2pq$	p^2	1

B$(2,\ p)$ の平均値は $2p$ であるから，分散の定義から

$$\begin{aligned}\sigma^2&=(0-2p)^2q^2+(1-2p)^22pq+(2-2p)^2p^2\\&=4p^2q^2+2pq-8p^2q+8p^3q\\&\quad+4p^2-8p^3+4p^4\\&=2p(2pq^2+q-4pq+4p^2q+2p\\&\quad-4p^2+2p^3)\\&=2p\{2(1-q)q^2+q-4(1-q)q+4(1-q)^2q\\&\quad+2(1-q)-4(1-q)^2+2(1-q)^3\}\\&=2p(2q^2-2q^3+q-4q+4q^2+4q\\&\quad-8q^2+4q^3+2-2q-4+8q-4q^2\\&\quad+2-6q+6q^2-2q^3)\\&=2pq\end{aligned}$$

(注) もう少し計算をうまくやることもできますが，ともかく，まともにやればめんどうだということがわかりましょう。

いよいよ，B$(n,\ p)$ の $\sigma^2=npq$ の証明です。二項定理から

$$(px+q)^n=\sum_{k=0}^{n}{}_n\mathrm{C}_kp^kq^{n-k}x^k$$

が成り立ちます。両辺を x で微分しますと

$$np(px+q)^{n-1}=\sum_{k=0}^{n}k{}_n\mathrm{C}_kp^kq^{n-k}x^{k-1}$$

この両辺に x を掛けますと

$$npx(px+q)^{n-1}=\sum_{k=0}^{n}k{}_n\mathrm{C}_kp^kq^{n-k}x^k$$

またも両辺を x で微分すれば

$$np\{(px+q)^{n-1}+(n-1)px(px+q)^{n-2}\}$$
$$=\sum_{k=0}^{n}k^2{}_n\mathrm{C}_kp^kq^{n-k}x^{k-1}$$

ここで $x=1$ とおきますと，$p+q=1$ より

$$np\{1+(n-1)p\}=\sum_{k=0}^{n}k^2{}_n\mathrm{C}_kp^kq^{n-k}$$

$$\therefore\quad \sigma^2=\sum_{k=0}^{n}k^2{}_n\mathrm{C}_kp^kq^{n-k}-(np)^2$$
$$=npq$$

Q.E.D.

*　*　*

◆ では，具体的な練習にいきましょう。

練習2. 二項分布 B$\left(10,\ \dfrac{1}{4}\right)$ の分散 σ^2 と標準偏差 σ を求めよ。

(ヒント) $\sigma^2=npq=10\cdot\dfrac{1}{4}\cdot\dfrac{3}{4}=\dfrac{15}{8}$

また，$\sigma=\sqrt{npq}=\sqrt{\dfrac{15}{8}}=\dfrac{\sqrt{30}}{4}$

練習3. 甲，乙2人がサイコロを振り，出た目が一致すれば甲の勝ちとする。90回勝負して，そのうち甲の勝った回数を X とするとき，X の分散を求めよ。

(ヒント) 甲と乙の出た目が一致する確率は

$\frac{6}{36}=\frac{1}{6}$ です。そして甲の勝った回数を確率変数Xとすると

$$P(X=x)={}_{90}C_x\left(\frac{1}{6}\right)^x\left(\frac{5}{6}\right)^{90-x}$$

となります。つまりXは≪二項分布 $B\left(90,\ \frac{1}{6}\right)$ に従う≫のです。

ゆえに，分散は

$$npq=90\cdot\frac{1}{6}\cdot\frac{5}{6}=\frac{25}{2}$$

となります。

■練習 4. 50本のくじの中に3本の当たりくじがある。4回引いたとき当たる回数をXとすれば，Xの標準偏差を求めよ。ただし，引いたくじはもとに戻すものとする。

(ヒント) くじを引いて当たる確率は $\frac{3}{50}$ で，4回引いたときX回当たる確率は

$${}_4C_X\left(\frac{3}{50}\right)^X\left(\frac{47}{50}\right)^{4-X}$$

ですから，Xは≪二項分布 $B\left(4,\ \frac{3}{50}\right)$ に従う≫のです。したがって，求める標準偏差は

$$\sigma=\sqrt{4\times\frac{3}{50}\times\frac{47}{50}}=\frac{\sqrt{141}}{25}$$

です。

*　*　*

◆ では，やや，総合的な問題をやってみましょう。

■練習 5. 1つのさいころを10回振るとき，1の目が出る回数を表す確率変数をXとする。

(1) $X=k$ $(0\leqq k\leqq 10)$ となる確率 $P(X=k)$ を求めよ。

(2) Xの期待値を求めよ。

(3) $\sum_{k=0}^{10}k^2P(X=k)$ の値を求めよ。

（新潟大）

(ヒント) (1) とくに，めんどうはないでしょう。ともあれさいころを振って1の目が出る確率は $\frac{1}{6}$ ですから，10回振ってk回1の目が出る確率は ${}_{10}C_k\left(\frac{1}{6}\right)^k\left(\frac{5}{6}\right)^{10-k}$ で与えられます。

(2) つまり，Xは二項分布 $B\left(10,\ \frac{1}{6}\right)$ に従うのです。したがって，Xの期待値は

$$E(X)=np=10\times\frac{1}{6}=\frac{5}{3}$$

です。

(3) また，分散は

$$\sigma^2=npq=10\cdot\frac{1}{6}\cdot\frac{5}{6}=\frac{25}{18}$$

となります。ところが，一方，分散は

$$\sigma^2=\sum_{k=0}^{10}k^2P(X=k)-\{E(X)\}^2$$

ですから

$$\sum_{k=0}^{10}k^2P(X=k)=\sigma^2+\{E(X)\}^2$$

$$=\frac{25}{18}+\left(\frac{5}{3}\right)^2=\frac{25}{6}$$

となります。

■練習 6. 二項分布 $P(k)={}_nC_kp^kq^{n-k}$ $(p+q=1,\ 0<p<1,\ k=0,\ 1,\ 2,\ \cdots,\ n)$ に対して $f(x)=\sum_{k=0}^{n}(x-k)^2P(k)$ とおくとき，$f(x)$ の最小値をnとpを用いて表せ。

（神奈川大）

(ヒント) $f(x)=\sum_{k=0}^{n}(x-k)^2P(k)$

$$=E\{(x-k)^2\}$$
$$=E\{x^2\}-E\{2xk\}+E\{k^2\}$$
$$=x^2-2xE\{k\}+E\{k^2\}$$
$$=\{x-E(k)\}^2+E\{k^2\}-\{E(k)\}^2$$

ゆえに，求める最小値は

$$E\{k^2\}-\{E(k)\}^2$$
$$=\sigma^2=npq=np(1-p)$$

となります。

*　*　*

◆ ともかく，うまいことをやろうと思わず定義に徹してやる覚悟が大切です。

（確率変数に対する）チェビシェフの不等式

1回目　年　月　日
2回目　年　月　日
3回目　年　月　日

◆チェビシェフという名前からどこの国の人と感じますか。彼はロシアの生んだすぐれた数学者であった。

◆ **チェビシェフの不等式**というのは，資料の扱いのところにも出ました（☞ p.126）。次は，確率変数の場合です。さて，それは：——

確率変数Xの平均をm，標準偏差をσとするとき，任意の正の数kに対して次の不等式が成り立つ。

$$P(|X-m|<k\sigma)\geqq 1-\frac{1}{k^2}$$

証明は次のようです。

$$X=\{x_1,\ x_2,\ \cdots\cdots,\ x_n\}$$

の順序を変えて，$|X-m|\geqq k\sigma$ を満足するものを

$$x_1',\ x_2',\ \cdots\cdots,\ x_i'$$

とし，$|X-m|<k\sigma$ を満足するものを

$$x_{i+1}',\ x_{i+2}',\ \cdots\cdots,\ x_n'$$

とし，さらに，上の各値に対応する確率をp_k' としますと

$$P(|X-m|<k\sigma)=p_{i+1}'+p_{i+2}'+\cdots+p_n'$$
$$=1-(p_1'+p_2'+\cdots\cdots+p_i')$$

という関係が成り立ちます。そして，

$$\sigma^2=(x_1'-m)^2p_1'+(x_2'-m)^2p_2'+\cdots\cdots$$
$$\cdots\cdots+(x_n'-m)^2p_n'$$
$$\geqq(x_1'-m)^2p_1'+(x_2'-m)^2p_2'+\cdots\cdots$$
$$\cdots\cdots+(x_i'-m)^2p_i'$$

ところが

$$(x_1'-m)^2\geqq k^2\sigma^2,\ (x_2'-m)^2\geqq k^2\sigma^2,\ \cdots\cdots$$
$$\cdots\cdots,\ (x_i'-m)^2\geqq k^2\sigma^2$$

ですから

$$\sigma^2\geqq k^2\sigma^2(p_1'+p_2'+\cdots\cdots+p_i')$$
$$\therefore\ \ \frac{1}{k^2}\geqq p_1'+p_2'+\cdots\cdots+p_i'$$
$$\therefore\ \ 1-\frac{1}{k^2}\leqq 1-(p_1'+p_2'+\cdots\cdots+p_i')$$
$$\therefore\ \ P(|X-m|<k\sigma)\geqq 1-\frac{1}{k^2}$$

Q.E.D.

＊　＊　＊

◆ このチェビシェフの不等式を使うと，次の**大数の法則**が証明できます。

ある試行において，事象Aが起こる確率をpとする。この試行を独立にn回くり返すとき，事象Aの起こる回数をXとすれば，任意の小さい正の数にα対して，次の式が成り立つ。

$$\lim_{n\to\infty}P\left(\left|\frac{X}{n}-p\right|<\alpha\right)=1$$

このことは重要です。つまり，nが十分大きいと，相対度数 $\frac{X}{n}$ は数学的確率pにほとんど等しくなるということを示すからです。

ところで，その証明には二項分布がものをいいます。すなわち：——

Xは二項分布に従うので，

$$m=np,\ \sigma=\sqrt{npq}$$

となります。これをチェビシェフの不等式に代入しますと

$$P(|X-np|<k\sqrt{npq})\geqq 1-\frac{1}{k^2}$$
$$\therefore\ \ P\left(\left|\frac{X}{n}-p\right|<k\sqrt{\frac{pq}{n}}\right)\geqq 1-\frac{1}{k^2}$$

そこで $k\sqrt{\frac{pq}{n}}=\alpha$ とおいて，上の式は

$$P\left(\left|\frac{X}{n}-p\right|<\alpha\right)\geqq 1-\frac{pq}{n\alpha^2}\ \cdots\cdots(*)$$

となります。ところが，p，q，αは定数ですから

$$\lim_{n\to\infty}P\left(\left|\frac{X}{n}-p\right|<\alpha\right)=1$$

となるのです。

◆ では，チェビシェフの不等式の応用をやってみませんか。

練習 1. 平均が 5，分散が 2 の確率分布で，変数 X が区間 $(3,\ 7)$ に属する確率は $\frac{1}{2}$ 以上であることを示せ。

ヒント　$m=5,\ \sigma=\sqrt{2}$ ですから

$3<X<7$ を変形しますと

$$-2<X-5<2$$

したがって

$$|X-5|<2=\sqrt{2}\cdot\sqrt{2}$$

これをチェビシェフの不等式とくらべて

$$k=\sqrt{2}$$

$$\therefore\quad P(|X-5|<\sqrt{2}\cdot\sqrt{2})\geqq 1-\frac{1}{(\sqrt{2})^2}$$
$$=\frac{1}{2}$$

となります。

練習 2. 高校3年の男子288人について英語の試験を行ったところ，平均点は62点で，標準偏差は6点であった。チェビシェフの不等式を用いて，50点以上で74点以下の生徒数は何人より多いといえるか。

ヒント　$(50-62)^2=(74-62)^2=12^2$
$$\leqq k^2\cdot 6^2$$

から

$$k^2\geqq 4$$

$$\therefore\quad \frac{n}{288}>1-\frac{1}{k^2}\geqq 1-\frac{1}{4}=\frac{3}{4}$$

$$\therefore\quad n>\frac{3}{4}\times 288=216$$

答　216人より多い

練習 3. 確率変数 X の平均が50，標準偏差が14であるとき，確率 $P(22\leqq X\leqq 78)$ の範囲を求めよ。

ヒント　平均 $m=50$，標準偏差 $\sigma=14$ ですから，チェビシェフの不等式において $k=2$ として

$$P(|X-50|\leqq 2\cdot 14)$$
$$=P(50-28\leqq X\leqq 50+28)$$
$$=P(22\leqq X\leqq 78)>1-\frac{1}{2^2}=\frac{3}{4}$$

ゆえに，$P(22\leqq X\leqq 78)$ は $\frac{3}{4}$ より大であることがわかります。

練習 4. 正しいさいころを n 回投げて1の目の出る回数を X とする。$\frac{X}{n}$ と $\frac{1}{6}$ の差が0.0005以下となる確率を0.9999以上にするためには n を何回以上とすればよいか。

ヒント　146ページの(*)をもういちど書くと，

$$P\left(\left|\frac{X}{n}-p\right|<\alpha\right)\geqq 1-\frac{pq}{n\alpha^2}$$

です。ここで，

$$\alpha=0.0005,\ p=\frac{1}{6},\ q=\frac{5}{6}$$

$$1-\frac{pq}{n\alpha^2}\geqq 0.9999$$

が成り立つようにすればよいのです。

すなわち

$$0.0001\geqq\frac{\frac{1}{6}\cdot\frac{5}{6}}{n\times 0.0005^2}$$

$$\therefore\quad n\geqq\frac{\frac{5}{36}}{0.0005^2\times 0.0001}$$
$$=\frac{10^{11}}{18}=\frac{100}{18}\times 10^9$$

したがって，$n>5555555555$

とすればよいのです。

注　この場合 $p=\frac{1}{6}$ ですが p の値がわからないときは $pq=p(1-p)=-\left(p-\frac{1}{2}\right)^2+\frac{1}{4}\leqq\frac{1}{4}$ ですから $\frac{5}{36}$ の代りに $\frac{1}{4}$ を入れて計算すると9999999999となります。

*　*　*

◆ 上のように $n>5555555555$ という答えにはいささか違和感を禁じえませんね。いっそ 5600000000 とでもしたらいいのに，と思うかも知れません。しかし，今のところは≪数学的に余りに数学的に≫です。

○二項分布表

1回目　年　月　日
2回目　年　月　日
3回目　年　月　日

◆二項分布はよく出てくるものですが，計算は単純ではありますが，労力は大きい。そこでこれを表にしておくと便利でしょう。

◆ 二項分布 B$(n,\ p)$ は，確率変数Xがr $(0\leqq r\leqq n)$ である確率が

$$P(X=r)={}_n\mathrm{C}_r p^r q^{n-r}\quad (q=1-p)$$

で与えられます。そこで，pのいろいろな値について ${}_n\mathrm{C}_r p^r q^{n-r}$ を計算しておくと便利でしょう。それが**二項分布表**です。

たとえば，B$\left(n,\ \dfrac{1}{2}\right)$の表は $n=10,\ 20,$ ……，100 について，各rの値に対する ${}_n\mathrm{C}_r p^r q^{n-r}$ の値を示してあります。

■練習 1. ${}_{50}\mathrm{C}_{20}\left(\dfrac{1}{2}\right)^{20}\left(\dfrac{1}{2}\right)^{50-20}$ の値を求めよ。

ヒント　B$\left(n,\ \dfrac{1}{2}\right)$ の表の $n=50$ の列について，$r=20$ のところをみると0.0419とあります。これが求める値です。

*　*　*

■練習 2. 硬貨を20回投げるとき，表の出る回数の平均を二項分布表を使って求めよ。

ヒント　硬貨を20回投げるとき，表の出る回数Xは二項分布 B$\left(20,\ \dfrac{1}{2}\right)$ に従います。そこで，表から次のようにして求めることができます。

r	$P(X=r)$	rp
0	0.0000	0.0000
1	0.0000	0.0000
2	0.0002	0.0004
3	0.0011	0.0033
4	0.0046	0.0184
5	0.0148	0.0740
6	0.0370	0.2220
7	0.0739	0.5173
8	0.1201	0.9608
9	0.1602	1.4418
10	0.1762	1.7620
11	0.1602	1.7622
12	0.1201	1.4412
13	0.0739	0.9607
14	0.0370	0.5180
15	0.0148	0.2220
16	0.0046	0.0736
17	0.0011	0.0187
18	0.0002	0.0036
19	0.0000	0.0000
20	0.0000	0.0000
計	1.0000	10.0000

もちろんながら，この結果は

$$m=np=20\cdot\frac{1}{2}=10$$

からすぐ求められるのですが，こうして計算してみて，はじめて気がつくことも多いものです。

■練習 3. 確率変数Xが B$(100,\ 0.1)$ に従うとき次の確率を求めよ。

$$P(10\leqq X\leqq 15)$$

ヒント　$\displaystyle\sum_{k=10}^{15}{}_{100}\mathrm{C}_k\left(\frac{1}{10}\right)^k\left(\frac{9}{10}\right)^{100-k}$

$=0.1319+0.1199+0.0988+0.0743$
$\quad+0.0513+0.0327$
$=0.5089$

これが正確な値です。

余裕があれば，正規近似（☞ p.162）による値を求めて上の結果とくらべてみるといいでしょう。

*　*　*

◆ 実際の入試では，二項分布表は部分的に与えられていることが多いものです。また正規近似でやることも多いのです。しかし，勉強としてはこの二項分布表を活用したいもの。

$\mathrm{B}\left(n, \frac{1}{10}\right) \quad P(X=r) = {}_n\mathrm{C}_r\left(\frac{1}{10}\right)^r\left(\frac{9}{10}\right)^{n-r} \quad \left(0 < r < \frac{n}{2},\ r\text{は整数}\right)$

r \ n	5	10	15	20	25	30	35	40	45	50
0	.5904 900	.3486 784	.2058 911	.1215 767	.0717 898	.0423 912	.0250 316	.0147 809	.0087 280	.0051 538
1	.3280 500	.3874 205	.3431 519	.2701 703	.1994 161	.1413 038	.0973 449	.0656 928	.0436 398	.0286 321
2	.0729 000	.1937 103	.2668 959	.2851 798	.2658 881	.2276 562	.1838 738	.1423 344	.1066 751	.0779 429
3		.0573 956	.1285 055	.1901 199	.2264 974	.2360 880	.2247 346	.2003 226	.1698 900	.1385 651
4		.0111 603	.0428 351	.0897 788	.1384 150	.1770 659	.1997 641	.2058 870	.1982 050	.1809 045
5			.0104 708	.0319 214	.0645 937	.1023 048	.1376 152	.1647 096	.1805 867	.1849 246
6			.0019 391	.0088 670	.0239 235	.0473 633	.0764 529	.1067 563	.1337 680	.1541 038
7			.0002 770	.0019 705	.0072 151	.0180 432	.0351 927	.0576 145	.0828 088	.1076 281
8				.0003 557	.0018 038	.0057 638	.0136 860	.0264 066	.0437 046	.0642 779
9				.0000 527	.0003 785	.0015 654	.0045 620	.0104 322	.0199 638	.0333 293
10					.0000 673	.0003 653	.0013 179	.0035 934	.0079 856	.0151 833
11					.0000 102	.0000 738	.0003 328	.0010 889	.0028 231	.0061 347
12					.0000 013	.0000 130	.0000 740	.0002 924	.0008 888	.0022 153
13						.0000 020	.0000 145	.0000 699	.0002 507	.0007 195
14						.0000 003	.0000 025	.0000 150	.0000 636	.0002 113
15							.0000 004	.0000 029	.0000 147	.0000 563
16							.0000 000	.0000 005	.0000 030	.0000 137
17							.0000 000	.0000 000	.0000 006	.0000 030
18								.0000 000	.0000 001	.0000 007
19								.0000 000	.0000 000	.0000 001
20									.0000 000	.0000 000
21									.0000 000	.0000 000
22									.0000 000	.0000 000
23										.0000 000
24										.0000 000

$P(0 \leqq X \leqq r) = \sum_{k=0}^{r} {}_n\mathrm{C}_k\left(\frac{1}{10}\right)^k\left(\frac{9}{10}\right)^{n-k} \quad \left(0 < r < \frac{n}{2},\ r\text{は整数}\right)$

r \ n	50	45	40	35	30	25	20	15	10	5
24	1.0000 000									
23	1.0000 000									
22	1.0000 000	1.0000 000								
21	1.0000 000	1.0000 000								
20	1.0000 000	1.0000 000								
19	1.0000 000	1.0000 000	1.0000 000							
18	.9999 999	1.0000 000	1.0000 000							
17	.9999 992	.9999 999	1.0000 000	1.0000 000						
16	.9999 962	.9999 993	.9999 999	1.0000 000						
15	.9999 825	.9999 963	.9999 994	.9999 999						
14	.9999 262	.9999 816	.9999 965	.9999 995	1.0000 000					
13	.9997 149	.9999 180	.9999 815	.9999 970	.9999 997					
12	.9989 954	.9996 673	.9999 116	.9999 825	.9999 977	.9999 998				
11	.9967 801	.9987 785	.9996 192	.9999 085	.9999 847	.9999 985				
10	.9906 454	.9959 554	.9985 303	.9995 757	.9999 109	.9999 883				
9	.9754 621	.9879 698	.9949 369	.9982 578	.9995 456	.9999 210	.9999 928			
8	.9421 328	.9680 060	.9845 047	.9936 958	.9979 802	.9995 425	.9999 401			
7	.8778 549	.9243 014	.9580 981	.9800 098	.9922 164	.9977 387	.9995 844	.9999 664		
6	.7702 268	.8414 926	.9004 836	.9448 171	.9741 732	.9905 236	.9976 139	.9996 894		
5	.6161 230	.7077 246	.7937 273	.8683 642	.9268 099	.9666 001	.9887 469	.9977 503		
4	.4311 984	.5271 379	.6290 177	.7307 490	.8245 051	.9020 064	.9568 255	.9872 795	.9983 651	
3	.2502 939	.3289 329	.4231 307	.5309 849	.6474 392	.7635 914	.8670 467	.9444 444	.9872 048	
2	.1117 288	.1590 429	.2228 081	.3062 503	.4113 512	.5370 940	.6769 268	.8159 389	.9298 092	.9914 400
1	.0337 859	.0523 678	.0804 737	.1223 765	.1836 950	.2712 059	.3917 470	.5490 430	.7360 989	.9185 400
0	.0051 538	.0087 280	.0147 809	.0250 316	.0423 912	.0717 898	.1215 767	.2058 911	.3486 784	.5904 900

$\mathrm{B}\left(n, \frac{1}{5}\right)$　$P(X=r)={}_{n}\mathrm{C}_{r}\left(\frac{1}{5}\right)^{r}\left(\frac{4}{5}\right)^{n-r}$　$\left(0<r<\frac{n}{2}\right.$, nは整数)

r ＼ n	5	10	15	20	25	30	35	40	45	50	
0	.3276 800	.1073 742	.0351 844	.0115 292	.0037 779	.0012 379	.0004 056	.0001 329	.0000 436	.0000 143	0
1	.4096 000	.2684 354	.1319 414	.0576 461	.0236 118	.0092 846	.0035 495	.0013 293	.0004 900	.0001 784	1
2	.2048 000	.3019 899	.2308 974	.1369 094	.0708 355	.0436 565	.0150 850	.0064 799	.0026 950	.0010 927	2
		.2013 266	.2501 389	.2053 642	.1357 681	.0785 318	.0414 839	.0205 200	.0096 572	.0043 710	3
24	.9999 979	.0880 804	.1876 042	.2181 994	.1866 810	.1325 225	.0829 677	.0474 524	.0253 502	.0128 396	4
23	.9999 917		.1031 823	.1745 595	.1960 151	.1722 791	.1286 000	.0854 143	.0519 679	.0295 312	5
22	.9999 698	.9999 967	.0429 926	.1090 997	.1633 459	.1794 575	.1607 500	.1245 626	.0866 131	.0553 710	6
21	.9998 978	.9999 865	.0138 191	.0545 498	.1108 419	.1538 207	.1664 910	.1512 545	.1206 397	.0870 116	7
20	.9996 793	.9999 492		.0221 609	.0623 486	.1105 586	.1456 797	.1559 812	.1432 597	.1169 218	8
19	.9990 676	.9998 236	.9999 783	.0073 870	.0294 423	.0675 637	.1092 597	.1386 500	.1472 391	.1364 088	9
18	.9974 888	.9994 372	.9999 148		.0117 770	.0354 708	.0710 189	.1074 538	.1325 152	.1398 190	10
17	.9937 392	.9983 496	.9996 955	.9999 658	.0040 148	.0161 232	.0403 516	.0732 639	.1054 098	.1271 082	11
16	.9855 583	.9955 530	.9990 090	.9998 586	.0011 711	.0063 821	.0201 758	.0442 636	.0746 653	.1032 754	12
15	.9691 966	.9889 955	.9970 638	.9994 752		.0022 091	.0089 239	.0238 342	.0473 837	.0754 705	13
14	.9392 779	.9750 060	.9920 841	.9982 481	.9997 688	.0006 707	.0035 058	.0114 915	.0270 765	.0498 644	14
13	.8894 135	.9479 295	.9805 926	.9947 423	.9990 981		.0012 271	.0049 797	.0139 895	.0299 187	15
12	.8139 430	.9005 458	.9567 584	.9858 184	.9968 890	.9996 310	.0003 834	.0019 452	.0065 575	.0163 617	16
11	.7106 676	.8258 805	.9124 948	.9656 426	.9905 069	.9984 599	.0001 072	.0006 865	.0027 966	.0081 809	17
10	.5835 594	.7204 707	.8392 309	.9252 910	.9743 837	.9944 451		.0002 193	.0010 876	.0037 496	18
9	.4437 404	.5879 555	.7317 771	.8542 721	.9389 129	.9826 681	.9974 052	.0000 635	.0003 864	.0015 788	19
8	.3073 316	.4407 164	.5931 271	.7450 124	.8713 492	.9532 258	.9900 182		.0001 256	.0006 117	20
7	.1904 098	.2974 567	.4371 459	.5993 327	.7607 906	.8908 772	.9678 573	.9957 603	.0000 373	.0002 185	21
6	.1033 982	.1768 170	.2858 914	.4328 417	.6069 699	.7800 353	.9133 075	.9819 412	.0000 102	.0000 720	22
5	.0480 272	.0902 039	.1613 288	.2720 917	.4275 124	.6166 894	.8042 078	.9389 486		.0000 219	23
4	.0184 960	.0382 360	.0759 145	.1434 917	.2552 333	.4206 743	.6296 483	.8357 663	.9672 065	.0000 082	24
3	.0056 564	.0128 858	.0284 621	.0605 240	.1227 108	.2339 933	.4114 489	.6481 621	.8791 261		
2	.0012 854	.0032 286	.0079 421	.0190 401	.0441 790	.0982 252	.2060 847	.3980 232	.6777 995	.9420 800	2
1	.0001 927	.0005 336	.0014 622	.0039 551	.0105 225	.0273 897	.0691 753	.1671 258	.3758 096	.7372 800	1
0	.0000 143	.0000 436	.0001 329	.0004 056	.0012 379	.0037 779	.0115 292	.0351 844	.1073 742	.3276 800	0
	50	45	40	35	30	25	20	15	10	5	n ＼ r

$P(0 \leqq X \leqq r)=\sum_{k=0}^{r}{}_{n}\mathrm{C}_{k}\left(\frac{1}{5}\right)^{k}\left(\frac{4}{5}\right)^{n-k}$　$\left(0<r<\frac{n}{2}\right.$, rは整数)

$$\mathrm{B}\left(n, \frac{3}{10}\right) \quad P(X=r) = {}_n\mathrm{C}_r\left(\frac{3}{10}\right)^r\left(\frac{7}{10}\right)^{n-r} \quad \left(0<r<\frac{n}{2},\ r\text{は整数}\right)$$

r \ n	5	10	15	20	25	30	35	40	45	50	
0	.1680 700	.0282 475	.0047 476	.0007 979	.0001 341	.0000 225	.0000 038	.0000 006	.0000 001	.0000 000	0
1	.3601 500	.1210 608	.0305 200	.0068 394	.0014 369	.0002 898	.0000 568	.0000 110	.0000 021	.0000 000	1
2	.3087 000	.2334 745	.0915 601	.0278 458	.0073 895	.0018 009	.0004 140	.0000 912	.0000 194	.0000 041	2
		.2668 279	.1700 402	.0716 037	.0242 800	.0072 034	.0019 517	.0004 951	.0001 196	.0000 277	3
24	.9976 305	.2001 210	.2186 232	.1304 210	.0572 314	.0208 383	.0066 915	.0019 630	.0005 378	.0001 397	4
23	.9944 078		.2061 303	.1788 630	.1030 165	.0464 399	.0177 803	.0060 572	.0018 902	.0005 510	5
22	.9877 239	.9975 754	.1472 360	.1916 390	.1471 665	.0829 282	.0381 006	.0151 428	.0054 007	.0017 708	6
21	.9749 130	.9940 397	.0811 301	.1642 620	.1711 936	.1218 537	.0676 480	.0315 220	.0128 954	.0047 705	7
20	.9522 362	.9864 773		.1143 967	.1650 796	.1501 412	.1014 721	.0557 263	.0262 514	.0109 891	8
19	.9151 974	.9716 549	.9937 455	.0653 696	.1336 359	.1572 908	.1304 641	.0849 162	.0462 524	.0219 783	9
18	.8594 401	.9450 506	.9852 230		.0916 360	.1415 617	.1453 742	.1128 173	.0713 609	.0386 190	10
17	.7821 931	.9013 671	.9680 487	.9935 815	.0535 535	.1103 078	.1415 983	.1318 644	.0973 102	.0601 855	11
16	.6838 786	.8358 418	.9366 871	.9840 392	.0267 768	.0748 517	.1213 700	.1365 738	.1181 625	.0838 297	12
15	.5691 784	.7462 153	.8848 534	.9641 178		.0444 176	.0920 278	.1260 682	.1285 504	.1050 174	13
14	.4468 316	.6346 800	.8074 482	.9269 310	.9830 627	.0231 152	.0619 778	.1041 991	.1259 269	.1189 484	14
13	.3278 832	.5087 531	.7032 491	.8649 532	.9599 475		.0371 868	.0774 052	.1115 353	.1223 468	15
12	.2228 658	.3802 027	.5771 809	.7729 254	.9155 299	.9825 303	.0199 214	.0518 337	.0896 265	.1147 002	16
11	.1390 361	.2620 402	.4406 071	.6515 554	.8406 782	.9557 535	.0095 423	.0313 616	.0655 253	.0983 145	17
10	.0788 506	.1647 300	.3087 427	.5099 571	.7303 704	.9022 000		.0171 743	.0436 835	.0772 470	18
9	.0402 316	.0933 691	.1959 254	.3645 829	.5888 087	.8105 640	.9520 381	.0085 225	.0266 043	.0557 573	19
8	.0182 533	.0471 167	.1110 092	.2341 188	.4315 179	.6769 281	.8866 685		.0148 224	.0370 388	20
7	.0072 642	.0208 653	.0552 829	.1326 467	.2813 767	.5118 485	.7722 718	.9499 875	.0075 624	.0226 768	21
6	.0024 937	.0079 699	.0237 609	.0649 987	.1595 230	.3406 549	.6080 098	.8688 574	.0035 357	.0128 109	22
5	.0007 229	.0025 692	.0086 181	.0268 981	.0765 948	.1934 884	.4163 708	.7216 214		.0066 839	23
4	.0001 719	.0006 790	.0025 609	.0091 178	.0301 549	.0904 719	.2375 078	.5154 911	.8497 317	.0032 227	24
3	.0000 322	.0001 412	.0005 979	.0024 263	.0093 166	.0332 405	.1070 868	.2968 679	.6496 107		
2	.0000 045	.0000 216	.0001 028	.0004 746	.0021 132	.0089 605	.0354 831	.1268 277	.3827 828	.8369 200	2
1	.0000 004	.0000 022	.0000 116	.0000 606	.0003 123	.0015 710	.0076 373	.0352 676	.1493 083	.5282 200	1
0	.0000 000	.0000 001	.0000 006	.0000 038	.0000 225	.0001 341	.0007 979	.0047 476	.0282 475	.1680 700	0
	50	45	40	35	30	25	20	15	10	5	n \ r

$$P(0 \leqq X \leqq r) = \sum_{k=0}^{r} {}_n\mathrm{C}_k\left(\frac{3}{10}\right)^k\left(\frac{7}{10}\right)^{n-k} \quad \left(0<r<\frac{n}{2},\ r\text{は整数}\right)$$

$\mathrm{B}\left(n, \dfrac{1}{4}\right) \quad P(X=r) = {}_n\mathrm{C}_r \left(\dfrac{1}{4}\right)^r \left(\dfrac{3}{4}\right)^{n-r} \quad \left(0 < r < \dfrac{n}{2},\ r\text{は整数}\right)$

r \ n	5	10	15	20	25	30	35	40	45	50
0	.2373 047	.0563 135	.0133 635	.0031 712	.0007 525	.0001 786	.0000 424	.0000 101	.0000 024	.0000 006
1	.3955 078	.1877 117	.0668 173	.0211 414	.0062 712	.0017 858	.0004 944	.0001 340	.0000 358	.0000 094
2	.2636 719	.2815 676	.1559 070	.0669 478	.0250 848	.0086 315	.0028 017	.0008 716	.0002 625	.0000 771
3		.2502 823	.2251 991	.1338 956	.0641 056	.0268 534	.0102 728	.0036 800	.0012 542	.0004 111
4		.1459 980	.2251 990	.1896 855	.1175 268	.0604 203	.0273 942	.0113 465	.0043 898	.0016 102
5			.1651 460	.2023 312	.1645 376	.1047 285	.0566 147	.0272 318	.0119 988	.0049 378
6			.0917 478	.1686 092	.1828 196	.1454 562	.0943 579	.0529 506	.0266 639	.0123 447
7			.0393 205	.1124 062	.1654 081	.1662 357	.1303 037	.0857 296	.0495 186	.0258 649
8				.0608 867	.1240 561	.1593 092	.1520 209	.1178 781	.0784 046	.0463 415
9				.0270 608	.0781 094	.1298 074	.1520 210	.1397 074	.1074 432	.0720 866
10					.0416 584	.0908 653	.1317 515	.1443 644	.1289 320	.0985 184
11					.0189 356	.0550 698	.0998 118	.1312 403	.1367 460	.1194 163
12					.0073 639	.0290 647	.0665 411	.1057 214	.1291 490	.1293 676
13						.0134 144	.0392 423	.0759 025	.1092 799	.1260 505
14						.0054 297	.0205 554	.0487 945	.0832 609	.1110 444
15							.0095 926	.0281 923	.0573 575	.0888 356
16							.0039 969	.0146 835	.0358 484	.0647 760
17							.0014 890	.0069 099	.0203 844	.0431 839
18								.0029 431	.0105 697	.0263 902
19								.0011 360	.0050 067	.0148 156
20									.0021 696	.0076 547
21									.0008 609	.0036 451
22									.0003 131	.0016 016
23										.0006 500
24										.0002 437

r \ n	50	45	40	35	30	25	20	15	10	5
24	.9998 775									
23	.9996 338									
22	.9989 838	.9998 519								
21	.9973 822	.9995 388								
20	.9937 371	.9986 779								
19	.9860 824	.9965 083	.9994 276							
18	.9712 668	.9915 016	.9982 916							
17	.9448 766	.9809 319	.9953 485	.9993 043						
16	.9016 927	.9605 475	.9884 386	.9978 153						
15	.8369 167	.9246 991	.9737 551	.9938 184						
14	.7480 811	.8673 416	.9455 628	.9842 258	.9972 505					
13	.6370 367	.7840 807	.8967 683	.9636 704	.9918 208					
12	.5109 862	.6748 008	.8208 658	.9244 281	.9784 064	.9966 296				
11	.3816 186	.5456 518	.7151 444	.8578 870	.9493 417	.9892 657				
10	.2622 023	.4089 058	.5839 041	.7580 752	.8942 719	.9703 301				
9	.1636 839	.2799 738	.4395 397	.6263 237	.8034 066	.9286 717	.9861 356			
8	.0915 973	.1725 306	.2998 323	.4743 027	.6735 992	.8505 623	.9590 748			
7	.0452 558	.0941 260	.1819 542	.3222 818	.5142 900	.7265 062	.8981 881	.9827 002		
6	.0193 909	.0446 074	.0962 246	.1919 781	.3480 543	.5610 981	.7857 819	.9433 797		
5	.0070 462	.0179 435	.0432 740	.0976 202	.2025 981	.3782 785	.6171 727	.8516 319		
4	.0021 084	.0059 447	.0160 422	.0410 055	.0978 696	.2137 409	.4148 415	.6864 859	9218 731	
3	.0004 982	.0015 549	.0046 957	.0136 113	.0374 493	.0962 141	.2251 560	.4612 869	.7758 751	
2	.0000 871	.0003 007	.0010 157	.0033 385	.0105 959	.0321 085	.0912 604	.2360 878	.5255 928	.8964 844
1	.0000 100	.0000 382	.0001 441	.0005 368	.0019 644	.0070 237	.0243 126	.0801 808	.2440 252	.6328 125
0	.0000 006	.0000 024	.0000 101	.0000 424	.0001 786	.0007 525	.0031 712	.0133 635	.0563 135	.2373 047

$P(0 \leqq X \leqq r) = \displaystyle\sum_{k=0}^{r} {}_n\mathrm{C}_k \left(\dfrac{1}{4}\right)^k \left(\dfrac{3}{4}\right)^{n-k} \quad \left(0 < r < \dfrac{n}{2},\ r\text{は整数}\right)$

$$\mathrm{B}\left(n, \frac{1}{2}\right) \quad P(X=r) = {}_n\mathrm{C}_r\left(\frac{1}{2}\right)^r\left(\frac{1}{2}\right)^{n-r} \quad \left(0 < r < \frac{n}{2},\ r\text{は整数}\right)$$

r \ n	5	10	15	20	25	30	35	40	45	50	
0	.0312 500	.0009 766	.0000 305	.0000 010	.0000 000	.0000 000	.0000 000	.0000 000	.0000 000	.0000 000	0
1	.1562 500	.0097 656	.0004 578	.0000 190	.0000 000	.0000 000	.0000 000	.0000 000	.0000 000	.0000 000	1
2	.3125 000	.0439 453	.0032 043	.0001 812	.0000 089	.0000 000	.0000 000	.0000 000	.0000 000	.0000 000	2
		.1171 875	.0138 855	.0010 872	.0000 686	.0000 038	.0000 000	.0000 000	.0000 000	.0000 000	3
24	.4438 624	.2050 781	.0416 565	.0046 206	.0003 770	.0000 255	.0000 015	.0000 000	.0000 000	.0000 000	4
23	.3359 055		.0916 443	.0147 857	.0015 834	.0001 328	.0000 095	.0000 006	.0000 000	.0000 000	5
22	.2399 438	.5000 000	.1527 405	.0369 644	.0052 779	.0005 530	.0000 472	.0000 035	.0000 000	.0000 000	6
21	.1611 182	.3829 959	.1963 806	.0739 289	.0143 260	.0018 959	.0001 957	.0000 169	.0000 013	.0000 000	7
20	.1013 194	.2757 422		.1201 343	.0322 335	.0054 510	.0006 850	.0000 700	.0000 061	.0000 005	8
19	.0594 602	.1856 490	.4373 147	.1601 792	.0608 854	.0133 246	.0020 550	.0002 487	.0000 252	.0000 022	9
18	.0324 543	.1163 466	.3179 140		.0974 166	.0279 816	.0053 428	.0007 709	.0000 906	.0000 091	10
17	.0164 196	.0675 782	.2147 953	.5000 000	.1328 409	.0508 756	.0121 429	.0021 026	.0002 885	.0000 332	11
16	.0076 733	.0362 271	.1340 936	.3679 394	.1549 810	.0805 531	.0242 857	.0050 812	.0008 174	.0001 078	12
15	.0033 002	.0178 489	.0769 300	.2497 799		.1115 351	.0429 671	.0109 442	.0020 750	.0003 152	13
14	.0013 011	.0080 472	.0403 452	.1552 523	.4277 678	.1354 354	.0675 197	.0211 065	.0047 428	.0008 330	14
13	.0004 681	.0033 044	.0192 387	.0877 326	.2923 324		.0945 276	.0365 848	.0098 017	.0019 991	15
12	.0001 529	.0012 294	.0082 945	.0447 655	.1807 973	.5000 000	.1181 595	.0571 636	.0183 782	.0043 731	16
11	.0000 451	.0004 120	.0032 133	.0204 798	.1002 442	.3450 190	.1320 606	.0807 017	.0313 511	.0087 463	17
10	.0000 119	.0001 235	.0011 107	.0083 369	.0493 686	.2121 781		.1031 187	.0487 684	.0160 347	18
9	.0000 028	.0000 329	.0003 398	.0029 941	.0213 870	.1147 615	.4119 015	.1194 007	.0693 024	.0270 059	19
8	.0000 006	.0000 077	.0000 911	.0009 391	.0080 624	.0538 761	.2517 223		.0900 932	.0418 592	20
7	.0000 001	.0000 016	.0000 211	.0002 541	.0026 114	.0216 426	.1315 880	.5000 000	.1072 537	.0597 988	21
6	.0000 000	.0000 003	.0000 042	.0000 584	.0007 155	.0073 166	.0576 591	.3036 194	.1170 041	.0788 256	22
5	.0000 000	.0000 000	.0000 007	.0000 112	.0001 625	.0020 387	.0206 947	.1508 789		.0959 617	23
4	.0000 000	.0000 000	.0000 001	.0000 017	.0000 297	.0004 553	.0059 090	.0592 346	.3769 531	.1079 569	24
3	.0000 000	.0000 000	.0000 000	.0000 002	.0000 042	.0000 783	.0012 884	.0175 781	.1718 750		
2	.0000 000	.0000 000	.0000 000	.0000 000	.0000 004	.0000 097	.0002 012	.0036 926	.0546 875	.5000 000	2
1	.0000 000	.0000 000	.0000 000	.0000 000	.0000 000	.0000 008	.0000 200	.0004 883	.0107 422	.1875 000	1
0	.0000 000	.0000 000	.0000 000	.0000 000	.0000 000	.0000 000	.0000 010	.0000 305	.0009 766	.0312 500	0
	50	45	40	35	30	25	20	15	10	5	n \ r

$$P(0 \leqq X \leqq r) = \sum_{k=0}^{r} {}_n\mathrm{C}_k\left(\frac{1}{2}\right)^k\left(\frac{1}{2}\right)^{n-k} \quad \left(0 < r < \frac{n}{2},\ r\text{は整数}\right)$$

◐ 正規分布とは何か

1回目　年　月　日
2回目　年　月　日
3回目　年　月　日

◈正規分布はわれわれの周囲に氾濫している。それなのに使う段になると，手がつかぬ，というのは，氾濫に流されているからです。

◆ 確率密度関数が

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

である確率分布を **正規分布** といいます。そして，ここに σ^2 は **分散** で，m は **平均** です。この分布を $N(m,\ \sigma^2)$ で表します。

特に，$m=0$, $\sigma=1$ のとき，すなわち，$N(0,\ 1)$ を標準正規分布といいます。この場合には，確率変数 x は

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

で表される分布になるわけです。そして，

$$\frac{1}{\sqrt{2\pi}}\int_0^{\lambda}e^{-\frac{x^2}{2}}dx=I(\lambda)$$

とおくと　$P\{m-\lambda\sigma\leqq x\leqq m+\lambda\sigma\}=2I(\lambda)$

となります。この $I(\lambda)$ を確率積分といい，詳しい数値表が作られています。簡略表を下にあげておきます。

≪確率積分表≫

λ	$I(\lambda)$	λ	$I(\lambda)$
0.1	0.0398	1.6	0.4452
0.2	0.0793	1.7	0.4554
0.3	0.1179	1.8	0.4641
0.4	0.1554	1.9	0.4713
0.5	0.1915	2.0	0.4772
0.6	0.2257	2.1	0.4821
0.7	0.2580	2.2	0.4861
0.8	0.2881	2.3	0.4893
0.9	0.3159	2.4	0.4918
1.0	0.3413	2.5	0.4938
1.1	0.3643	2.6	0.4953
1.2	0.3849	2.7	0,4965
1.3	0.4032	2.8	0.4974
1.4	0.4192	2.9	0.4981
1.5	0.4332	3.0	0.4987

◆ では，この表を用いて次の練習をやってみませんか。

■練習 1. 変量 X が $m-\sigma$ と $m+\sigma$ の間にある確率を求めよ。

(ヒント) $\lambda=1$ ですから，表をみると 0.3413 になっています。2倍すると，0.6826 です。つまり，変量 X が $m-\sigma$ と $m+\sigma$ の間にある確率は 0.6826 に等しいのです。

■練習 2. 変量 X が $m-2\sigma$ と $m+2\sigma$ の間にある確率を求めよ。

(ヒント) $\lambda=2$ ですから，表をみると 0.4772 です。2倍すると 0.9544 です。したがって変量 X が $m-2\sigma$ と $m+2\sigma$ の間にある確率は 0.9544 であることがわかります。

■練習 3. $N(10,\ 4)$ において $P\{X\geqq 13\}$ を求めよ。

(ヒント) $N(10,\ 4)$ というのは平均値10，分散4の正規分布ですが，左の表を使うためには変数を変換しなければなりません。

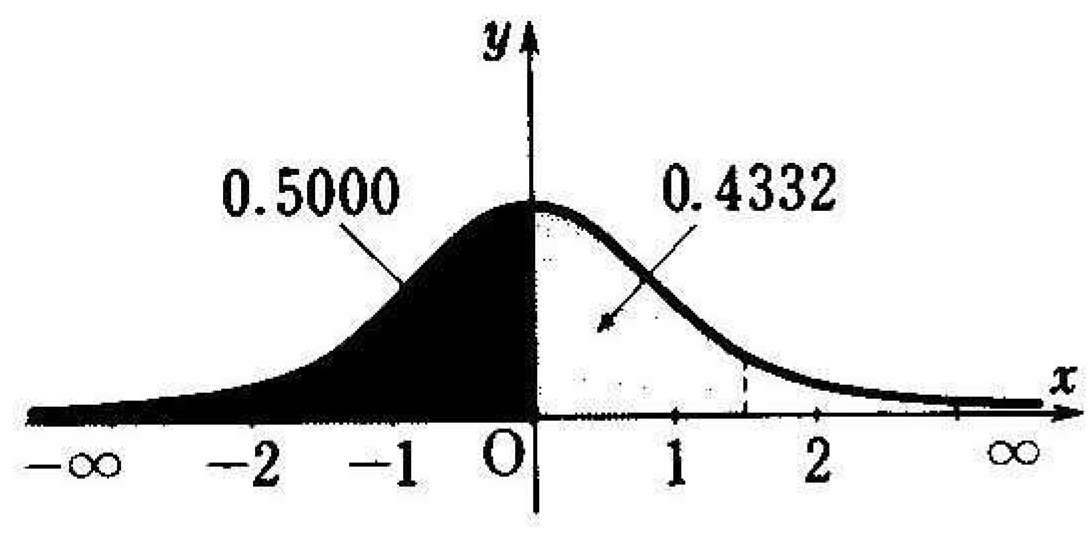
$\dfrac{x-m}{\sigma}=t$ とおきますと

$$P\{X\geqq 13\}=\int_{\frac{13-10}{2}}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt$$

$$=1-\int_{-\infty}^{\frac{3}{2}}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt$$

$$=1-(0.5000+0.4332)$$

$$=1-0.9332=0.0668$$



◆ では，いよいよ統計らしい問題をやってみませんか。

練習 4. クラスの者の英語と数学の成績を調べたら次の表のごとくであった。K君の成績は英語76点，数学70点であった。クラスの中の成績順位からいうと，K君は英語と数学とではどちらが上位にあると考えられるか。理由をつけて答えなさい。（慶大）

	英　語	数　学
平 均 点	69 点	56 点
標準偏差	8.2 点	12.4点

（ヒント） この場合には **標準測度**（ひょうじゅんそくど）というものを使います。つまり，

$$\frac{x-m}{\sigma}$$

なる量で，これは平均からの差 $x-m$ が標準偏差 σ の何倍であるかを示すもの。

この場合

英語では　$\dfrac{76-69}{8.2}=0.854$

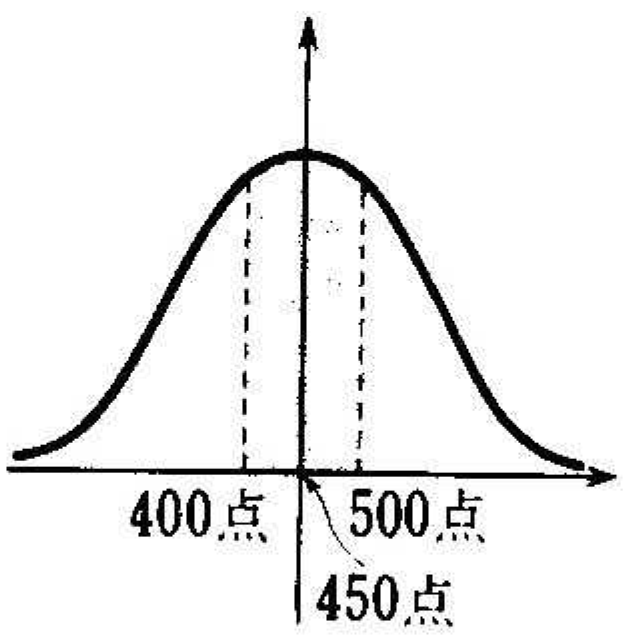
数学では　$\dfrac{70-56}{12.4}=1.129$

ですから，数学のほうが上位であることがわかります。

練習 5. ある大学の入学試験において，志望者の得点は平均450点，標準偏差100点であった。得点400点以上500点未満の得点者は何パーセントあったと考えられるか。ただし，志望者の得点は正規分布をなすものとする。

（解） 志望者の成績 x が平均450，標準偏差100の正規分布をなすから

$$t=\frac{x-m}{\sigma}=\frac{x-450}{100}$$



の分布は，正規分布 $N(0,\ 1)$ となる。ゆえに400点以上500点未満の人数の割合は

$$P(400 \leqq t < 500)$$
$$=P\left(\frac{400-450}{100} \leqq t < \frac{500-450}{100}\right)$$
$$=P(-0.5 \leqq t < 0.5)=0.1915\times 2$$
$$=0.3830$$

となるから，約38％あったと考えられる。

（注） 上の解で 0.1915 はもちろん確率積分表が与えられたとして求めたものです。

練習 6. 方程式 $ax^2+4bx+c=0$ が相異なる2つの実数解をもつ確率を次の(1)，(2)のそれぞれの場合に対して求めよ。

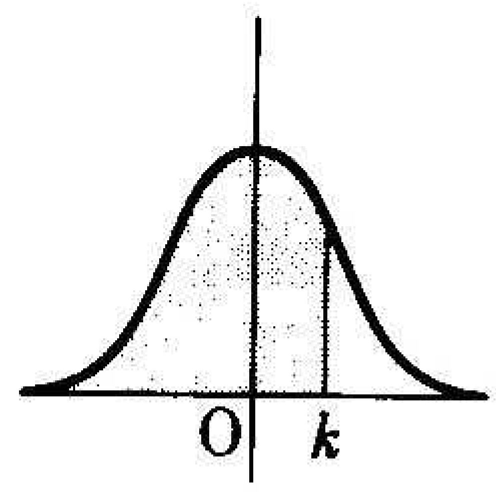
(1)　$a,\ b,\ c$ がそれぞれ無作為に 0，1，2 のいずれかの値をとる場合。

(2)　$a,\ c$ がそれぞれ無作為に 1，2 のいずれかの値をとり，$a,\ c$ と関係なく b は平均0，標準偏差1の正規分布にしたがう場合。

ただし，$\sqrt{2}=1.4$ とし，次の正規分布表を利用せよ。

ここに，

$$S(k)=\frac{1}{\sqrt{2\pi}}\int_0^k e^{-\frac{x^2}{2}}dx$$



は図の陰影部分を表す。

k	0.3	0.5	0.7	0.9	1.0	1.2
$S(k)$	0.12	0.19	0.26	0.32	0.34	0.38

（静岡大）

（ヒント） $ax^2+4bx+c=0$ が相異なる2つの実数解をもつ条件は

$$a \neq 0,\ 4b^2-ac>0$$

(1)　全部で $3^3=27$（通り）ありますが，アタリは全部目の子勘定で調べてみても11しかありませんから，求める確率は $\dfrac{11}{27}$ です。

(2)　$a,\ c$ のとり方は $2^2=4$（通り）あって，$ac=1$ となるのは $a=1,\ c=1$ の1通りで，このとき $b>0.5,\ b<-0.5$ ですが，b がこの範囲の値をとる確率は，正規分布表から $0.62(=1-0.19\times 2)$ となります。$ac=2,4$ のときも同様にして調べて答は 0.475 です。

⑩ 正規分布の平均値と標準偏差の理論

1回目　年　月　日
2回目　年　月　日
3回目　年　月　日

◆二項分布の方は理論的にそれであることがわかっているのに対し，正規分布の方は経験的にしかわかっていないことが多いのだ。

◆ Xの確率密度関数 $f(x)$ が

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

であるとき，Xは正規分布 $\mathrm{N}(m,\ \sigma^2)$ に従うといいます。

正規分布 $\mathrm{N}(m,\ \sigma^2)$ において

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

とおくと，Zの確率分布は

$$g(z)=\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}$$

で表される

これは，平均値0，標準偏差1の正規分布 $\mathrm{N}(0,\ 1)$ でこれを**標準正規分布**といいます。

*　*　*

◆ では，練習問題をやってみませんか。

■練習 1. 確率変数Xが平均10，標準偏差 2 の正規分布に従うとき，Xの確率密度関数を書け。

ヒント　$m=10,\ \sigma=2$ ですから

$$f(x)=\frac{1}{\sqrt{2\pi}\cdot 2}e^{-\frac{(x-10)^2}{2\cdot 2^2}}$$

すなわち

$$f(x)=\frac{1}{2\sqrt{2\pi}}e^{-\frac{(x-10)^2}{8}}$$

となります。

■練習 2. 確率密度関数が

$$f(x)=\frac{1}{2\sqrt{\pi}}e^{-\frac{(x-5)^2}{4}}$$

で与えられる正規分布を標準化せよ。

ヒント　$f(x)$ を変形すると

$$f(x)=\frac{1}{\sqrt{2\pi}\sqrt{2}}e^{-\frac{(x-5)^2}{2\cdot\sqrt{2}^2}}$$

となりますから

平均 $m=5$，標準偏差 $\sigma=\sqrt{2}$

です。したがって変数変換

$$Z=\frac{X-m}{\sigma}=\frac{X-5}{\sqrt{2}}$$

を行うと $\mathrm{N}(0,\ 1)$ となります。

■練習 3. Xは，平均値がm，標準偏差がσの正規分布に従う確率変数とする。

(1) 確率 $P(\alpha\leqq X\leqq\beta)$ を，関数 $\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$ の定積分の形で表せ。ただし，$\alpha<\beta$ とする。

(2) aを定数とするとき，次の不等式を満足する正数εの範囲を求めよ。

$$P(|X-m|\leqq\varepsilon)>\sqrt{\frac{2}{\pi}}\int_0^a e^{-\frac{x^2}{2}}dx$$

（岩手大）

ヒント　(1) Xが $\mathrm{N}(m,\ \sigma^2)$ に従うというのですから $\frac{X-m}{\sigma}$ は $\mathrm{N}(0,\ 1)$ に従います。

$$\therefore\quad P(\alpha\leqq X\leqq\beta)=\int_{\frac{\alpha-m}{\sigma}}^{\frac{\beta-m}{\sigma}}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx$$

となりましょう。

(2) $P(|X-m|\leqq\varepsilon)$

$=P(m-\varepsilon\leqq X\leqq m+\varepsilon)$

と表せます。そして，これは

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

ですから

$$2\int_0^{\frac{\varepsilon}{\sigma}}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx>\sqrt{\frac{2}{\pi}}\int_0^a e^{-\frac{x^2}{2}}dx$$

$$\therefore\quad \int_0^{\frac{\varepsilon}{\sigma}}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx>\int_0^a\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx$$

$$\therefore \quad \frac{\varepsilon}{\sigma}>a$$

$$\therefore \quad \varepsilon>a\sigma \qquad \cdots\cdots 答$$

*　*　*

◆ では，やや，総合的な問題をとりあげてみましょう。

■練習 4. 確率変数Xが平均80，標準偏差10の正規分布に従うとき，確率変数 X^2 の平均 $E(X^2)$ を求めよ。

ヒント　$\{\sigma(X)\}^2$ を算出する式を思い出して下さい。

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

でした。したがって

$$E(X^2)=\{\sigma(X)\}^2+\{E(X)\}^2$$
$$=10^2+80^2$$
$$=6500 \qquad \cdots\cdots 答$$

■練習 5. 確率変数Xが正規分布 $N(m,\ \sigma^2)$ に従うとき，$Z=\dfrac{X-m}{\sigma}$ も正規分布に従うことが分っている。このとき，Zの分布は $N(0,\ 1)$ であることを示せ。

ヒント　$Z=\dfrac{X-m}{\sigma}$

$$\therefore \quad E(Z)=E\left(\frac{X}{\sigma}-\frac{m}{\sigma}\right)=E\left(\frac{X}{\sigma}\right)-E\left(\frac{m}{\sigma}\right)$$
$$=\frac{1}{\sigma}E(X)-\frac{m}{\sigma}=\frac{m}{\sigma}-\frac{m}{\sigma}=0$$

また，$\{\sigma(X)\}^2=\left\{\sigma\left(\dfrac{X}{\sigma}-\dfrac{m}{\sigma}\right)\right\}^2$

$$=\left\{\sigma\left(\frac{X}{\sigma}\right)\right\}^2=\frac{1}{\sigma^2}\{\sigma(X)\}^2=\frac{\sigma^2}{\sigma^2}=1$$

$$\therefore \quad \{\sigma(X)\}^2=1$$

ゆえに，仮定「Zは正規分布に従う」より，Zは $N(0,\ 1)$ に従うことがわかります。

■練習 6. 平均値μ，分散 σ^2 の正規分布を $N(\mu,\ \sigma^2)$ とする。変数xが $N(1,\ 2^2)$ に従うとき，$1\leqq x\leqq 5$ の確率と，$N(0,\ 3^2)$ に従うとき，$0\leqq x\leqq a$ の確率とが等しい。aを求めよ。　　(自治医大)

ヒント　両方とも標準正規分布に直してくらべればよいでしょう。つまり

$$P(1\leqq x\leqq 5)=P\left(\frac{1-1}{2}\leqq z\leqq\frac{5-1}{2}\right)$$
$$=P(0\leqq z\leqq 2)$$

そして

$$P(0\leqq x\leqq a)=P\left(\frac{0-0}{3}\leqq z\leqq\frac{a-0}{3}\right)$$
$$=P\left(0\leqq z\leqq\frac{a}{3}\right)$$

この2つが等しいというのですから

$$\frac{a}{3}=2, \qquad \therefore \quad a=6$$

答　$a=6$

では，もうひとつ。

■練習 7. 変量Zの分布が標準正規分布 $N(0,\ 1)$ であるとき，確率 $\varphi(z)=P(0\leqq Z\leqq z)$ に対して次の表がある。

z	0	1	2	3
$\varphi(z)$	0	0.34134	0.47725	0.49865

この表を用いて，次の確率を求めよ。

(1) $P(1\leqq Z\leqq 3)$

(2) $P(-2\leqq Z\leqq 1)$

次に変量Xの分布が正規分布 $N(m,\ \sigma^2)$ であるとき

(3) $P(|X-m|\leqq 2\sigma)$ を求めよ。

ヒント　(1) $P(1\leqq Z\leqq 3)$

$=P(0\leqq Z\leqq 3)-P(0\leqq Z\leqq 1)$

$=\varphi(3)-\varphi(1)=0.49865-0.34134$

$=0.15731$

(2) $P(-2\leqq Z\leqq 1)$

$=P(-2\leqq Z\leqq 0)+P(0\leqq Z\leqq 1)$

$=P(0\leqq Z\leqq 2)+P(0\leqq Z\leqq 1)$

$=0.47725+0.34134$

$=0.81859$

(3) $P(|X-m|\leqq 2\sigma)=P\left(\dfrac{|X-m|}{\sigma}\leqq 2\right)$

$=P(-2\leqq Z\leqq 2)=2P(0\leqq Z\leqq 2)$

$=2\times 0.47725$

$=0.95450$

① 正規分布の平均値と標準偏差の応用

1回目　年　月　日
2回目　年　月　日
3回目　年　月　日

◆ここでは，正規分布の平均値と標準偏差について，具体的な問題を扱うことにしましょう。大切なことは理論の部で終わりました。

◆ では，さっそくながら，次の練習へいくとしようではないか。

■練習 1. ある市の高校生の身長を Xcm とし，Xは$N(165,\ 5^2)$ に従うものとする。このうち１人を任意に選んだとき，その生徒の身長が，170cm 以上である確率を求めよ。

ヒント　$Z=\dfrac{X-165}{5}$ とおきますと，Zは標準正規分布 $N(0,\ 1)$ に従うのですから，正規分布表を用いて

$$P(170\leqq X)=P\left(\frac{170-165}{5}\leqq Z\right)$$
$$=P(1\leqq Z)=0.5-P(0\leqq Z\leqq 1)$$
$$=0.5-0.3413=0.1587$$ ……答

■練習 2. ある高校３年生の男子500人の身長の分布は，ほぼ正規分布をなしている。平均値が169cm，標準偏差が5.5cmであるとき，身長が160cm以上170cm未満の生徒は何人ぐらいいるか。

ヒント　身長を Xcm としますと

$Z=\dfrac{X-169}{5.5}$ は標準正規分布$N(0,\ 1)$ に従います。そこで，

$$P(160\leqq X<170)$$
$$=P\left(\frac{160-169}{5.5}\leqq Z<\frac{170-169}{5.5}\right)$$
$$=P(-1.6\leqq Z<0.18)$$
$$=P(0\leqq Z\leqq 1.6)+P(0\leqq Z\leqq 0.18)$$

正規分布表を用いて

$$=0.4452+0.0714$$
$$=0.5166$$

よって，求める人数は

$$500\times 0.5166=258.3（人）$$

とすることができます。

■練習 3. ある大学の入学試験は1000点満点で，全志願者2000名の得点の分布は，平均点450，標準偏差75の正規分布をなしていることがわかった。また，入学定員は320名である。

以上のことから，合格最低点を求めよ。

（小樽商大）

ヒント　Xは正規分布（450，75^2）に従うから

$Z=\dfrac{X-m}{\sigma}=\dfrac{X-450}{75}$ は標準正規分布$N(0,1)$に従います。

そこで，合格最低点を a としますと

$$P(X\geqq a)=\frac{320}{2000}=0.16$$

となるハズ。したがって

$$P(m\leqq X<a)=0.5-0.16=0.34$$

ところが正規分布表から

$$P(0\leqq Z<1)=0.34$$

（正規分布表では0.34134ですが）

$$\therefore\ \frac{a-450}{75}=1$$
$$\therefore\ a=525（点）$$ ……答

＊　＊　＊

◆ ではやや総合的な問題にいくとしましょう。

■練習 4. (1)　微分方程式 $\dfrac{dy}{dt}=-ky$（kは定数）の解で，$t=0$ のとき $y=5$ となるものを求めよ。

(2)　(1)の解 y の $t=a$（$a>0$）における値が $y\leqq 1$ をみたすとき，k の最小値

$M(a)$ を求めよ。ただし，自然対数 $\log 5=1.6$ とする。

(3) $M(a)$ を(2)で求めたものとする。k が平均 1，分散 $\frac{1}{9}$ の正規分布 $N\left(1, \frac{1}{9}\right)$ に従うとき確率 $P(k \geqq M(1))$ と $P(k \geqq M(8))$ を求めよ。また，$P(k \geqq M(n)) \geqq 0.945$ となる最小の整数 n を求めよ。ただし，標準正規分布 N(0, 1) の確率 $\Phi(z)=\frac{1}{\sqrt{2\pi}}\int_0^z e^{-\frac{x^2}{2}}dx$ については次の数表を用いよ。

z	1.4	1.6	1.8	2.0	2.2	2.4
$\Phi(x)$	0.419	0.445	0.464	0.477	0.486	0.492

（佐賀大）

ヒント (1) 変数分離形の微分方程式です。

$$\int \frac{dy}{y} = -\int k\,dt$$

$$\therefore \quad \log|y| = -kt+C$$

$$\therefore \quad |y| = e^{-kt+C}$$

$$y = \pm e^C e^{-kt}$$

ここで改めて $\pm e^C = A$ と書くと

$$y = Ae^{-kt}$$

$t=0$ のとき $y=5$ ですから $A=5$

$$\therefore \quad y=5e^{-kt} \qquad \cdots\cdots \text{答}$$

(2) $t=a$ のとき $y \leqq 1$ ですから

$$5e^{-ka} \leqq 1$$

$$\therefore \quad \log 5 - ka \leqq 0$$

$$\therefore \quad k \geqq \frac{\log 5}{a} = \frac{1.6}{a}$$

$$\therefore \quad M(a) = \frac{8}{5a} \qquad \cdots\cdots \text{答}$$

(3) $M(a)=\frac{8}{5a}$ ですから

$$M(1)=\frac{8}{5}$$

$$\therefore \quad P(k \geqq M(1)) = P\left(k \geqq \frac{8}{5}\right)$$

ところが：——

確率変数が正規分布に従うときの確率を思い出して，次のようになります。

$$\frac{\frac{8}{5}-1}{\frac{1}{3}} = 1.8$$

で表から

$$0.5-0.464=0.036$$

次に，

$$P(k \geqq M(8)) = P\left(k \geqq \frac{1}{5}\right)$$

ところが

$$\frac{\frac{1}{5}-1}{\frac{1}{3}} = -2.4$$

で，表から

$$0.5+0.492=0.992$$

となります。最後に

$$P(k \geqq M(n)) \geqq 0.945$$
$$= 0.5+0.445$$

ですから，表により

$$\frac{\frac{8}{5n}-1}{\frac{1}{3}} \leqq -1.6$$

$$\therefore \quad 7n \geqq 24, \qquad \therefore \quad n \geqq 4$$

つまり，求める n の最小の整数は 4 であることがわかります。

では，最後にひとつ：——

■練習 5. 錠剤中の有効成分量が平均 200 単位，標準偏差 100 単位の正規分布に従っているとき，50～350単位の範囲内にある錠剤は全体の約何%を占めると考えられるか。 （静岡薬大）

ヒント X は N(200, 100^2) に従うというのですから，$Z=\frac{X-200}{100}$ とおいて標準化しますと

$$P(50 \leqq X \leqq 350)$$
$$= P(-1.5 \leqq Z \leqq 1.5) = 2P(0 \leqq Z \leqq 1.5)$$
$$= 2 \times 0.4332 = 0.8664$$

答 86.6%

べつにメンドウはありませんでしたね。

⑪ 標準測度の扱い方

1回目　年　月　日
2回目　年　月　日
3回目　年　月　日

◆英語の点数が60点，数学の点数が50点であるからといって，数学が不得意といってはいけませんよ。なぜなら……

◆ さっそくながら，具体的な問題にふれてみましょう。

■**練習 1.** あるクラスでは国語，数学の平均点はそれぞれ43.2点，52.3点で，標準偏差はそれぞれ5.3点，12.4点であった。このクラスのA君の成績は国語47点，数学55点であった。A君は国語と数学といずれが得意といってよいか。

(ヒント) 平均点を上まわり，あるいは下まわる分が標準偏差の何倍になるかで，その人の点が全体の中で占める順位をきめることができます。この場合であれば，A君について，国語の場合には

$$\frac{47-43.2}{5.3}\fallingdotseq 0.72$$

数学については

$$\frac{55-52.3}{12.4}\fallingdotseq 0.22$$

つまり，この場合，A君の点数は国語より数学の方がよいのですが，平均から上まわる割合でいえば国語の方がグッとはなれています。してみると，A君は，国語の方が数学よりも得意であるといってよいでしょう。

(注) 上のように

$$\frac{\text{変量}-\text{平均値}}{\text{標準偏差}}$$

の値を**標準測度**というのです。だから，実は標準測度は変数Xを標準化した確率変数つまりZなのですが，このコトバを使うときのニュアンスがちょっとちがうということです。

*　*　*

◆ では，次の例にいくとしましょう。いいですか，コトバにとられてはいけませんよ。言葉を変えて内容まで変ったとの誤解は怖い。

■**練習 2.** ある生徒の得点と各科目の平均点，標準偏差が右のようであった。この生徒の成績はどの科目が最上位にあるといえるか。

	得点	平均点	標準偏差
国語	70	56	12
数学	68	50	14
英語	74	62	10

(ヒント) 標準測度を求めてみますと

国語：$\dfrac{70-56}{12}\fallingdotseq 1.17$

数学：$\dfrac{68-50}{14}\fallingdotseq 1.29$

英語：$\dfrac{74-62}{10}\fallingdotseq 1.20$

これによると，数学の点はもっとも低いが，数学が最上位にあることがわかります。

*　*　*

◆ 今一歩進めて，順位を求める問題をやってみましょう。

■**練習 3.** 100人の生徒に，数学のテストをしたところ，その成績は，平均点が58点，標準偏差が14点の正規分布をなしていた。このテストで72点とった人は上位からだいたい何番目とみてよいか。

(ヒント) 標準正規分布 N(0, 1) に従う確率変数について標準測度をUとするとき

$$P(|U|<1)=0.6826$$
$$P(|U|<2)=0.9544$$
$$P(|U|<3)=0.9974$$

です。このだいたいの値をオボエていれば，正規分布表なしでも扱える場合があります。

この練習3. では，標準測度は

$$\frac{72-58}{14}=1$$

ですから，まさに，上の結果が使えます。

すなわち，

$$P(|U|\geqq 1)=0.5-\frac{0.6826}{2}=0.1587$$

したがって，上から15.87％のところにあることがわかります。そして，それは

$$100\times 0.1587=15.87$$

つまり，100人中，上から16番目くらいのところということがわかります。

練習 4. 1000人の生徒に数学のテストをしたところ，その成績は，平均点60点，標準偏差が15点の正規分布をなしていた。このテストで23番目の人はほぼ何点をとったと考えてよいか。

ヒント　23番目ということはそれより上に入る確率は $\frac{23}{1000}=0.023$ です。

ところが前ページ右の

$$P(|U|<2)=0.9544$$

から

$$P(U\geqq 2)=\frac{1}{2}(1-0.9544)$$

$$=0.0228$$

となりますから，この人の標準測度はほぼ2と考えてよいでしょう。だから

$$\frac{X-60}{15}=2$$

より

$$X=90$$

つまり，この人は90点ぐらいとったと考えてよいでしょう。

答　約90点

*　*　*

◆ 上の各練習問題は，確率分布表を使わない範囲でやりました。しかし，次には，堂々と使うことにしましょう。では，これです。

練習 5. あるテストをしたところ，平均点64点，標準偏差12点になった。上位5％の者は評価A，次の15％に評価Bを与えたい。評価A，Bの与えられる者の点数の幅を求めよ。

ヒント　点数を X としますと，$Z=\frac{X-64}{12}$ は N(0, 1) に従います。

今，a 点以上の人には A，$b\leqq X<a$ の人にはBが与えられるとしますと，

$$P(X\geqq a)=0.05$$

$$P(b\leqq X<a)=0.15$$

$$\therefore\quad P\left(0\leqq Z<\frac{a-64}{12}\right)=0.45$$

正規分布表（☞ p.164）から

$$\frac{a-64}{12}=1.65$$

4

$$\therefore\quad a=83.8$$

次に，

$$P(b\leqq X)-P(a\leqq X)=0.15$$

より

$$P(b\leqq X)=0.15+0.05=0.20$$

$$\therefore\quad P\left(0\leqq Z<\frac{b-64}{12}\right)=0.50-0.20=0.30$$

正規分布表から

$$\frac{b-64}{12}=0.84$$

$$\therefore\quad b=74.08$$

ゆえに，Aは84点以上，Bは74点以上84点未満とすればよいでしょう。

注　もちろん，このテストを受けた人が十分多く，それが正規分布をしていることは，アンアンリに仮定してあるわけです。

なお，5％で15％ではなく，2.3％と16.9％であれば，160ページ右の数値を使って，つまり，正規分布表を使わずにできるというわけです。

また，正規分布をしていないかも知れないときはチェビシェフの不等式（☞p.146）を使うより仕方がないことになります。

*　*　*

◆ このように確率変数に対する制限が大きくなるにつれ，扱い方はめんどうになるかわり，くわしい知識ないし情報が得られるのです。そして，その知識や情報は，その確率変数自体の中に含まれていることが多いのです。

⑪ 二項分布の正規近似

1回目　　年　　月　　日
2回目　　年　　月　　日
3回目　　年　　月　　日

◆二項分布 B(n, p) はnの値が大きいとき，正規分布で近似させることができます。その方法がここの目的です。

◆ 正規分布についてまだ学んでない人はもちろん，あとまわしにして下さい。ここでは正規分布はやってあるとして，やりますから。

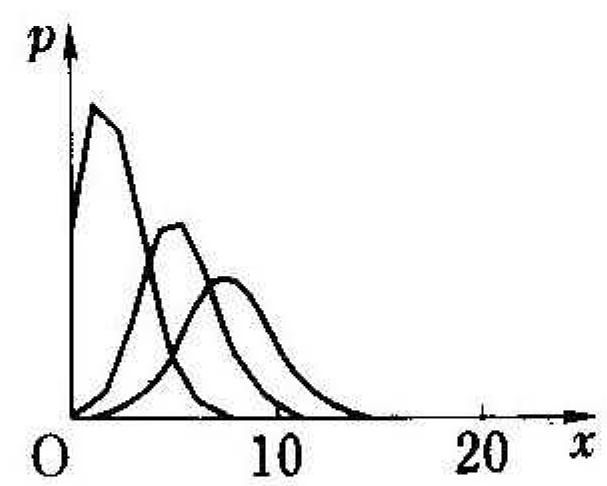
さて，それは：——

nが十分大きく，Xが B(n, p) に従うとき，平均値m，標準偏差σとして

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

は，標準正規分布 N(0, 1) に従うとみなすことができる

$p=\frac{1}{6}$ のときについて，$n=10$, 30, 50 とふえるに従って，${}_nC_xp^xq^{n-x}$ のグラフを図に示したのが右下のグラフです。nが大きくなるにつれて正規分布に近づいてゆくことがわかるでしょう。



さて：——

練習 1. 二項分布 B$\left(100, \frac{1}{2}\right)$ を正規近似せよ。

ヒント　$n=100$, $p=\frac{1}{2}$ ですから

$$m=np=50$$
$$\sigma=\sqrt{npq}=5$$

です。そして，

$$Z=\frac{X-50}{5}$$

は標準正規分布 N(0, 1) に従うのです。

——では，次も問題はありますまい。

練習 2. 二項分布B$\left(180, \frac{1}{6}\right)$ を正規近似せよ。

ヒント　$n=180$, $p=\frac{1}{6}$ ですから，

$$m=np=180\times\frac{1}{6}=30$$

$$\sigma=\sqrt{npq}=\sqrt{180\cdot\frac{1}{6}\cdot\frac{5}{6}}=5$$

で，したがって $Z=\frac{X-30}{5}$ は N(0, 1) に従うのです。

*　*　*

◆ では，具体的な問題にしきましょう。

練習 3. 確率変数Xが二項分布 B(100, 0.2) に従うとき $P(X=20)$ を求めよ。

ヒント　二項分布はいわゆる離散的な分布であるのに，正規分布は連続な分布です。そこで $X=20$ というのは

$$19.5\leqq X\leqq 20.5$$

に対応するものとします。(☞次頁注を見よ)

ところで，

$$m=np=100\times0.2=20$$

$$\sigma=\sqrt{npq}=\sqrt{100\times0.2\times0.8}=4$$

ですから

$$\frac{19.5-20}{4}=\frac{-0.5}{4}=-0.125$$

$$\frac{20.5-20}{4}=\frac{0.5}{4}=0.125$$

となり，　$19.5\leqq X\leqq 20.5$　は

$$-0.125\leqq Z\leqq 0.125$$

に対応します。かくして

$$P(X=20)\fallingdotseq P(-0.125\leqq Z\leqq 0.125)$$
$$\fallingdotseq P(-0.13\leqq Z\leqq 0.13)$$

正規分布表から0.13に対する値は

$$(0.04776+0.05567)\div 2\fallingdotseq 0.05172$$

ですから，求める確率は

$$P(X=20) \fallingdotseq 0.05172\times 2$$
$$\fallingdotseq 0.1034$$

となります。

■**練習 4.** 1枚の硬貨を100回投げるとき，表が60回以上出る確率を求めよ。

(ヒント) 表の出る回数をXとしますとXは二項分布$\mathrm{B}\left(100,\ \frac{1}{2}\right)$に従います。そして，

$$m=np=100\times\frac{1}{2}=50$$
$$\sigma=\sqrt{npq}=\sqrt{100\times\frac{1}{2}\times\frac{1}{2}}=5$$

ですから，

$$Z=\frac{X-m}{\sigma}=\frac{X-50}{5}$$

は正規分布 N(0, 1) をなすとみてよいのです。そして，$X=60$ に対応するZの値は

$$\frac{59.5-50}{5}=1.9$$

ですから

$$P(X\geqq 60)=P(Z\geqq 1.9)$$

正規分布表から

$$=0.5-0.47128$$
$$=0.02872$$

* * *

◆ では，やや総合的な練習にいきましょう。

■**練習 5.** さいころを3000回振るとき，1の目が460回以上540回以下出る確率を求めよ。

(ヒント) 1の目の出る回数をXとしますと，Xは二項分布 $\mathrm{B}\left(3000,\ \frac{1}{6}\right)$ に従います。

そして，

平均値 $m=3000\times\frac{1}{6}=500$

標準偏差 $\sigma=\sqrt{3000\times\frac{1}{6}\times\frac{5}{6}}=\sqrt{\frac{1250}{3}}$

$$\fallingdotseq 20.4$$

となります。

そこで

$$Z=\frac{X-500}{20.4}$$

とおくとZは正規分布 N(0, 1) に従うとしてよいのです。かくして

$$P(460\leqq X\leqq 540)$$
$$\fallingdotseq P\left(\frac{459.5-500}{20.4}\leqq Z\leqq\frac{540.5-500}{20.4}\right)$$
$$\fallingdotseq P(-1.985\leqq Z\leqq 1.985)$$
$$\fallingdotseq(-1.99\leqq Z\leqq 1.99)$$

正規分布表から

$$\fallingdotseq\frac{0.47615+0.47725}{2}\times 2$$
$$=0.95340\fallingdotseq 0.95$$

(注) この場合は $n=3000$ が十分大きいので0.5の補正（これを**半整数補正**といいます）をしなくても同じ結果が得られます。

■**練習 6.** ある工場に3つの食堂があり，毎日昼食時に，これら食堂を同時に利用して食事をする従業員が200人いるものとする。各人は等しい確率で1つの食堂を選ぶものとする。利用者が95%以上の確率で座れるためには，それぞれの食堂で最低いくつの座席を用意すればよいか。 （静岡大）

(ヒント) 1つの食堂に入る人数をXとしますと，Xは二項分布$\mathrm{B}\left(200,\ \frac{1}{3}\right)$に従います。

したがって，

$$m=200\times\frac{1}{3}=\frac{200}{3}$$
$$\sigma=\sqrt{200\times\frac{1}{3}\times\frac{2}{3}}=\frac{20}{3}$$

で，$Z=\dfrac{X-\frac{200}{3}}{\frac{20}{3}}$ は N(0, 1) に従うのです。

かくして

$$P(Z\leqq x_0)=0.95$$

となる x_0 を求めればよいわけです。正規分布表を使って x_0 を求めて

$$X_0\geqq 77.6\cdots\cdots$$

が得られます。 答 78席以上

1回目　年　月　日
2回目　年　月　日
3回目　年　月　日

○正規分布表

◆統計を扱う限り正規分布から離れることはできないのです。それにしても，正規分布のもつこの妖しさは何だろう。

◆ 表の意味と使い方については（☞ p.154）をみてください。

$$\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}},\ I(x)=\int_0^x \varphi(x)dx$$

◆正規分布表(1)◆

x	$\varphi(x)$	$I(x)$	x	$\varphi(x)$	$I(x)$	x	$\varphi(x)$	$I(x)$
0.00	.39894	.00000	0.92	.26129	.32121	1.84	.07341	.46712
.02	.39886	.00798	.94	.25647	.32639	.86	.07074	.46856
.04	.39862	.01595	.96	.25164	.33147	.88	.06814	.46995
.06	.39822	.02392	.98	.24681	.33646	1.90	.06562	.47128
.08	.39767	.03188	1.00	.24197	.34134	.92	.06316	.47257
0.10	.39695	.03983	.02	.23713	.34614	.94	.06077	.47381
.12	.39608	.04776	.04	.23230	.35083	.96	.05844	.47500
.14	.39505	.05567	.06	.22747	.35543	.98	.05618	.47615
.16	.39387	.06356	.08	.22265	.35993	2.00	.05399	.47725
.18	.39253	.07142	1.10	.21785	.36433	.02	.05186	.47831
0.20	.39104	.07926	.12	.21307	.36864	.04	.04980	.47932
.22	.38940	.08706	.14	.20831	.37286	.06	.04780	.48030
.24	.38762	.09483	.16	.20357	.37698	.08	.04586	.48124
.26	.38568	.10257	.18	.19886	.38100	2.10	.04398	.48214
.28	.38361	.11026	1.20	.19419	.38493	.12	.04217	.48300
0.30	.38139	.11791	.22	.18954	.38877	.14	.04041	.48382
.32	.37903	.12552	.24	.18494	.39251	.16	.03871	.48461
.34	.37654	.13307	.26	.18037	.39617	.18	.03706	.48537
.36	.37391	.14058	.28	.17585	.39973	2.20	.03547	.48610
.38	.37115	.14803	1.30	.17137	.40320	.22	.03394	.48679
0.40	.36827	.15542	.32	.16694	.40658	.24	.03246	.48745
.42	.36526	.16276	.34	.16256	.40988	.26	.03103	.48809
.44	.36213	.17003	.36	.15822	.41309	.28	.02965	.48870
.46	.35889	.17724	.38	.15395	.41621	2.30	.02833	.48928
.48	.35553	.18439	1.40	.14973	.41924	.32	.02705	.48983
0.50	.35207	.19146	.42	.14556	.42220	.34	.02582	.49036
.52	.34849	.19847	.44	.14146	.42507	.36	.02463	.49086
.54	.34482	.20540	.46	.13742	.42786	.38	.02349	.49134
.56	.34105	.21226	.48	.13344	.43056	2.40	.02239	.49180
.58	.33718	.21904	1.50	.12952	.43319	.42	.02134	.49224
0.60	.33322	.22575	.52	.12566	.43574	.44	.02033	.49266
.62	.32918	.23237	.54	.12188	.43822	.46	.01936	.49305
.64	.32506	.23891	.56	.11816	.44062	.48	.01842	.49343
.66	.32086	.24537	.58	.11450	.44295	2.50	.01753	.49379
.68	.31659	.25175	1.60	.11092	.44520	.52	.01667	.49413
0.70	.31225	.25804	.62	.10741	.44738	.54	.01585	.49446
.72	.30785	.26424	.64	.10396	.44950	.56	.01506	.49477
.74	.30339	.27035	.66	.10059	.45154	.58	.01431	.49506
.76	.29887	.27637	.68	.09728	.45352	2.60	.01358	.49534
.78	.29431	.28230	1.70	.09405	.45543	.62	.01289	.49560
0.80	.28969	.28814	.72	.09089	.45728	.64	.01223	.49585
.82	.28504	.29389	.74	.08780	.45907	.66	.01160	.49609
.84	.28034	.29955	.76	.08478	.46080	.68	.01100	.49632
.86	.27562	.30511	.78	.08183	.46246	2.70	.01042	.49653
.88	.27086	.31057	1.80	.07895	.46407	.72	.00987	.49674
0.90	.26609	.31594	.82	.07614	.46562	.74	.00935	.49693

x	$\varphi(x)$	$I(x)$	x	$\varphi(x)$	$I(x)$	x	$\varphi(x)$	$I(x)$
2.76	.00885	.49711	3.32	.00161	.49955	3.88	.00021	.49995
.78	.00837	.49728	.34	.00151	.49958	3.90	.00020	.49995
2.80	.00792	.49744	.36	.00141	.49961	.92	.00018	.49996
.82	.00748	.49760	.38	.00132	.49964	.94	.00017	.49996
.84	.00707	.49774	3.40	.00123	.49966	.96	.00016	.49996
.86	.00668	.49788	.42	.00115	.49969	.98	.00014	.49997
.88	.00631	.49801	.44	.00107	.49971	4.00	.00013	.49997
2.90	.00595	.49813	.46	.00100	.49973	.04	.00011	.49997
.92	.00562	.49825	.48	.00094	.49975	.08	.00010	.49998
.94	.00530	.49836	3.50	.00087	.49977	.12	.00008	.49998
.96	.00499	.49846	.52	.00081	.49978	.16	.00007	.49998
.98	.00471	.49856	.54	.00076	.49980	4.20	.00006	.49999
3.00	.00443	.49865	.56	.00071	.49981	.24	.00005	.49999
.02	.00417	.49874	.58	.00066	.49983	.28	.00004	.49999
.04	.00393	.49882	3.60	.00061	.49984	.32	.00004	.49999
.06	.00370	.49889	.62	.00057	.49985	.36	.00003	.49999
.08	.00348	.49897	.64	.00053	.49986	4.40	.00002	.49999
3.10	.00327	.49903	.66	.00049	.49987	.44	.00002	.50000
.12	.00307	.49910	.68	.00046	.49988	.48	.00002	.50000
.14	.00288	.49916	3.70	.00042	.49989	.52	.00001	.50000
.16	.00271	.49921	.72	.00039	.49990	.56	.00001	.50000
.18	.00254	.49926	.74	.00037	.49991	4.60	.00001	.50000
3.20	.00238	.49931	.76	.00034	.49992	.64	.00001	.50000
.22	.00224	.49936	.78	.00031	.49992	.68	.00001	.50000
.24	.00210	.49940	3.80	.00029	.49993	.72	.00001	.50000
.26	.00196	.49944	.82	.00027	.49993	.76	.00001	.50000
.28	.00184	.49948	.84	.00025	.49994	∞	.00000	.50000
3.30	.00172	.49952	.86	.00023	.49994			

◆正規分布表(2)◆

$$I(x)=\frac{1}{\sqrt{2\pi}}\int_0^x e^{-\frac{x^2}{2}}dx \longrightarrow x$$

$I(x)$	x	$I(x)$	x	$I(x)$	x	$I(x)$	x	$I(x)$	x
0.000	.0000	0.100	.2533	0.200	.5244	0.300	.8416	0.400	1.282
.004	.0100	.104	.2637	.204	.5359	.304	.8560	.404	1.305
.008	.0201	.108	.2741	.208	.5476	.308	.8705	.408	1.329
.012	.0301	.112	.2845	.212	.5592	.312	.8853	.412	1.353
.016	.0401	.116	.2950	.216	.5710	.316	.9002	.416	1.379
0.020	.0502	0.120	.3055	0.220	.5828	0.320	.9154	0.420	1.405
.024	.0602	.124	.3160	.224	.5948	.324	.9307	.424	1.433
.028	.0702	.128	.3266	.228	.6068	.328	.9463	.428	1.461
.032	.0803	.132	.3372	.232	.6189	.332	.9621	.432	1.491
.036	.0904	.136	.3478	.236	.6311	.336	.9782	.436	1.522
0.040	.1004	0.140	.3585	0.240	.6433	0.340	.9945	0.440	1.555
.044	.1105	.144	.3692	.244	.6557	.344	1.011	.444	1.589
.048	.1206	.148	.3799	.248	.6682	.348	1.028	.448	1.626
.052	.1307	.152	.3907	.252	.6808	.352	1.045	.452	1.665
.056	.1408	.156	.4016	.256	.6935	.356	1.063	.456	1.706
0.060	.1510	0.160	.4125	0.260	.7063	0.360	1.080	0.460	1.751
.064	.1611	.164	.4234	.264	.7192	.364	1.098	.464	1.799
.068	.1713	.168	.4344	.268	.7323	.368	1.117	.468	1.852
.072	.1815	.172	.4454	.272	.7454	.372	1.136	.472	1.911
.076	.1917	.176	.4565	.276	.7588	.376	1.155	.476	1.977
0.080	.2019	0.180	.4677	0.280	.7722	0.380	1.175	0.480	2.054
.084	.2121	.184	.4789	.284	.7858	.384	1.195	.484	2.144
.088	.2224	.188	.4902	.288	.7995	.388	1.216	.488	2.257
.092	.2327	.192	.5015	.292	.8134	.392	1.237	.492	2.409
.096	.2430	.196	.5129	.296	.8274	.396	1.259	.496	2.652

$1-2I(x)$	.05	.01	.001	.0001
x	1.9600	2.5758	3.2905	3.8906

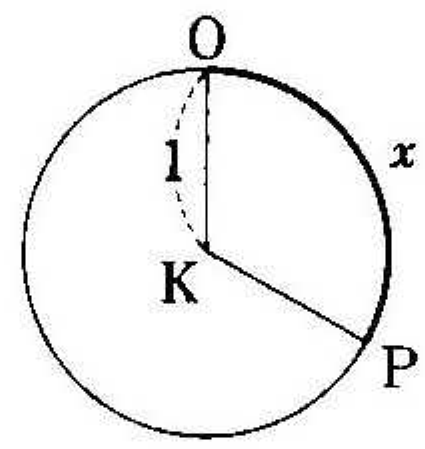
一様分布では

1回目　年　月　日
2回目　年　月　日
3回目　年　月　日

◆確率密度関数のグラフがx軸に平行なら，確率はいうまでもなく，一定です。これを一様分布ということがあります。

◆ まず，下のような場合を考えてみましょう。

練習 1. 半径1の円上に定点Oをとり，この点から，時計まわりに測って，円周上の任意の点にいたる弧の長さを確率変数Xとする。



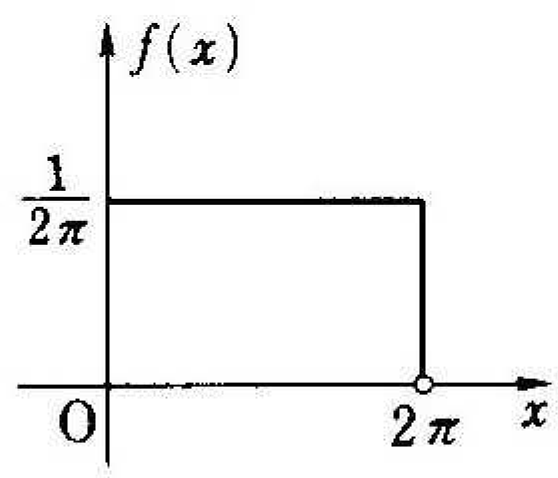
(1) 確率密度関数 $f(x)$ を求めよ。
(2) Xの平均を求めよ。
(3) Xの分散を求めよ。

ヒント $0 \leqq x < 2\pi$ であることはわかるでしょう。

(1) 確率密度関数 $f(x)=k$ (一定) ですから

$$\int_0^{2\pi} k\,dx = 2\pi k = 1$$

$$\therefore \quad k=\frac{1}{2\pi}$$



となります。つまり

$$f(x)=\begin{cases} 0 \leqq x < 2\pi : \dfrac{1}{2\pi} \\ 2\pi \leqq x : 0 \end{cases}$$

コノヨウニ，確率密度関数が定数のときこれを**一様分布**といいます。

(2) 平均mは定義によって

$$m=\int_0^{2\pi} xf(x)dx=\int_0^{2\pi} x\cdot\frac{1}{2\pi}dx$$

$$=\frac{1}{2\pi}\left[\frac{x^2}{2}\right]_0^{2\pi}=\pi$$

(3) 分散 σ^2 は定義によって

$$\sigma^2=\int_0^{2\pi}(x-m)^2 f(x)dx$$

$$=\int_0^{2\pi}(x-\pi)^2\frac{1}{2\pi}dx=\frac{1}{2\pi}\left[\frac{(x-\pi)^3}{3}\right]_0^{2\pi}$$

$$=\frac{1}{6\pi}\{(\pi)^3-(-\pi)^3\}$$

$$=\frac{\pi^2}{3}$$

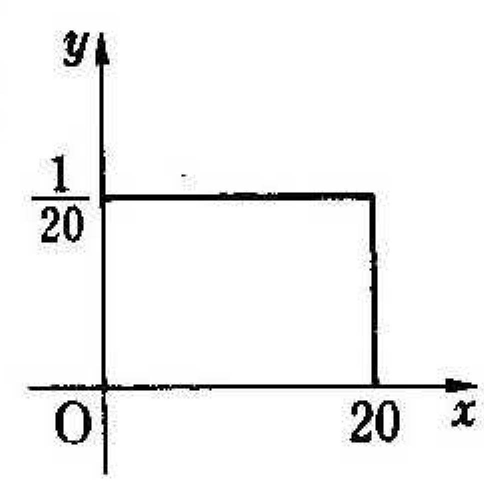
となります。

では，もうひとつ：――

練習 2. ちょうど20分おきに発車するバスに乗るのに，でたらめに停留所に行った場合の待ち時間を確率変数Xとする。

(1) 確率密度関数 $f(x)$ を求めよ。
(2) Xの平均を求めよ。
(3) Xの分散を求めよ。

ヒント (1) Xは区間$[0, 20]$のすべての値をとり，確率密度関数は



$$f(x)=k \text{ (一定)}$$

で与えられます。そして

$$\int_0^{20} f(x)dx=\int_0^{20} k\,dx=\Big[kx\Big]_0^{20}$$

$$=20k=1$$

$$\therefore \quad k=\frac{1}{20}$$

$$\therefore \quad f(x)=\frac{1}{20}$$

(2) Xの平均は定義によって

$$E(X)=\int_0^{20} xf(x)dx=\int_0^{20} x\cdot\frac{1}{20}dx$$

$$=\frac{1}{20}\left[\frac{x^2}{2}\right]_0^{20}=10 \text{ (分)}$$

(3) 分散 σ^2 は定義によって

$$\sigma^2=\int_0^{20}(x-10)^2\frac{1}{20}dx$$

$$=\frac{1}{20}\left[\frac{(x-10)^3}{3}\right]_0^{20}$$

$$=\frac{1}{60}(10^3-(-10)^3)=\frac{100}{3}$$

です。なお，待ち時間の標準偏差は

$$\sigma=\sqrt{\frac{100}{3}}=\frac{10}{\sqrt{3}}=\frac{10\sqrt{3}}{3}$$

となります。

＊　＊　＊

◆ では，やや，総合的な問題をやってみませんか。

練習 3. A駅発B駅行の電車には，特急，快速急行，急行の3種類があり，所要時間はそれぞれ35分，37分，40分である。特急は毎時0分，快速急行は毎時20分，40分，急行は毎時x分，$(x+20)$分，$(x+40)$分に発車する。(ただし，xは $1\leqq x\leqq 15$ である整数）乗客はA駅に着いた後，常に最初に発車する電車に乗るものとする。時刻表を知らずにA駅に来た乗客がa分発の電車に乗る確率は，その直前の電車がb分発だったとすると $\dfrac{a-b}{60}$ であり，その乗客がA駅に着いてから発車までの待ち時間の平均値（期待値）は $\dfrac{a-b}{2}$ 分である。(ただし，$a=0$ のときは，これら2数は $\dfrac{60-b}{60}$，$\dfrac{60-b}{2}$ でおきかえる）

(1) 乗客が，A駅についてからB駅に到着するまでの時間（待ち時間と電車の所要時間の和）の平均値 M をxで表せ。

(2) Mを最小にする整数xの値とMの最小値を求めよ。

（奈良女大）

ヒント (1) 乗客がA駅に着いた時刻を0時t分としますと，xは $1\leqq x\leqq 15$ の整数ですから，$0<t\leqq x$ のときには，x分発の急行に乗ることになりましょう。

このとき所要時間は $40+\dfrac{x-0}{2}$，x分発に乗る確率は $\dfrac{x-0}{60}$ となります。

同じように考えて次の表が得られます。

時刻(t)	$0<t\leqq x$	$x<t\leqq 20$	$20<t\leqq x+20$
時　間	$40+\frac{x}{2}$	$37+\frac{20-x}{2}$	$40+\frac{x}{2}$
確　率	$\frac{x}{60}$	$\frac{20-x}{60}$	$\frac{x}{60}$
時刻(t)	$x+20<t\leqq 40$	$40<t\leqq x+40$	$x+40<t\leqq 60$
時　間	$37+\frac{20-x}{2}$	$40+\frac{x}{2}$	$35+\frac{20-x}{2}$
確　率	$\frac{20-x}{60}$	$\frac{x}{60}$	$\frac{20-x}{60}$

そこで，

$$\begin{aligned}M=&\frac{x}{60}\left(\frac{80+x}{2}\right)+\frac{20-x}{60}\left(\frac{94-x}{2}\right)\\&+\frac{x}{60}\left(\frac{80+x}{2}\right)+\frac{20-x}{60}\left(\frac{94-x}{2}\right)\\&+\frac{x}{60}\left(\frac{80+x}{2}\right)+\frac{20-x}{60}\left(\frac{90-x}{2}\right)\\=&\frac{1}{120}\{3x(x+80)+2(x-20)(x-94)\\&+(x-20)(x-90)\}\\=&\frac{1}{60}(3x^2-49x+2780)\end{aligned}$$

(2) 上の式を変形して，平方を完成すれば

$$M=\frac{1}{20}\left(x-\frac{49}{6}\right)^2-\frac{49^2}{720}+\frac{139}{3}$$

ところがxは整数ですから，$\dfrac{49}{6}\fallingdotseq 8.2$ より，$x=8$ のときMは最小になって，その値は43です。

＊　＊　＊

◆ どうです，わかりますか。もし，ピンとこなかったら，上の表で示した関係を少し大きな方眼紙に丁寧に書いてみることです。定規でキチンと書くのだよ。

その上でもういちどはじめから考えなおしてみるといいでしょう。今ひとつの方法は，いわゆる鉄道グラフ状のものをかいてみるのです。耳なれない人は，物理でいう s-t グラフのようなものです。こうしてみると，それを別な観点からみることにもなるのです。

折れ線分布では

1回目　年　月　日
2回目　年　月　日
3回目　年　月　日

◆確率密度関数のグラフが線分でつながれているとき，つまり折れ線のとき，これを折れ線分布ということにしよう。

◆ たとえば，次のような問題をとりあげてみましょう。

練習 1. 区間$\left[-\dfrac{m}{2},\ \dfrac{n}{2}\right]$のすべての値をとる確率変数$X$の確率密度関数が

$$f(x)=\begin{cases}\dfrac{1}{m}x+\dfrac{1}{2} & \left(-\dfrac{m}{2}\leqq x\leqq 0\right)\\ -\dfrac{1}{n}x+\dfrac{1}{2} & \left(0\leqq x\leqq \dfrac{n}{2}\right)\end{cases}$$

によって与えられている。Xの平均値が$\dfrac{2}{3}$のとき，次のものを求めよ。

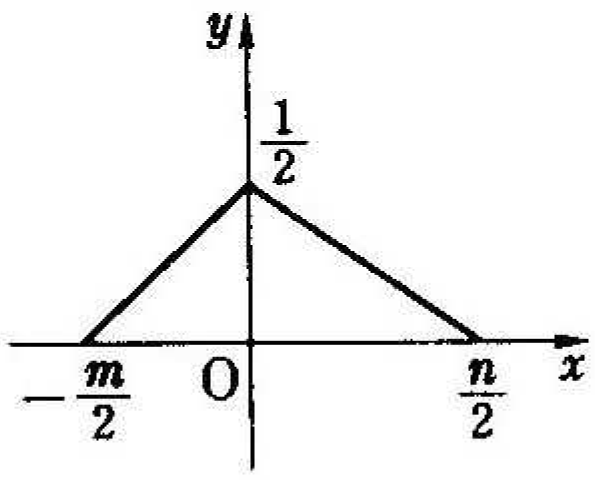
(1)　m，nの値

(2)　Xの分散

(3)　tの２次方程式　$t^2-4t+2X+1=0$ の２つの解がともに正である確率

（静岡大）

ヒント　(1)　確率密度関数のグラフは右のようになります。これとx軸を囲む面積が１にならなければなりません。



$$\therefore\ \int_{-\frac{m}{2}}^{\frac{n}{2}} f(x)dx=1$$

実は積分するまでもなく，三角形の面積から

$$\frac{1}{2}\left(\frac{m}{2}\right)\times\frac{1}{2}+\frac{1}{2}\left(\frac{n}{2}\right)\times\frac{1}{2}=1$$

$$\therefore\ m+n=8 \qquad \cdots\cdots①$$

となります。

次に，平均が$\dfrac{2}{3}$ですから，定義により

$$E(X)=\int_{-\frac{m}{2}}^{\frac{n}{2}} xf(x)dx$$

コレハ仕方ガナイ，積分スルトシヨウ。

$$=\int_{-\frac{m}{2}}^{0} xf(x)dx+\int_{0}^{\frac{n}{2}} xf(x)dx$$

$$=\int_{-\frac{m}{2}}^{0}\left(\frac{1}{m}x^2+\frac{x}{2}\right)dx+\int_{0}^{\frac{n}{2}}\left(-\frac{x^2}{n}+\frac{x}{2}\right)dx$$

$$=\frac{m^2}{24}-\frac{m^2}{16}-\frac{n^2}{24}+\frac{n^2}{16}=\frac{2}{3}$$

これより　　$n^2-m^2=32$　　……②

①，②を解くと

$$m=2,\ n=6$$

(2)　分散は定義により求めます。まず，

$$E(X^2)=\int_{-1}^{0} x^2\left(\frac{1}{2}x+\frac{1}{2}\right)dx+\int_{0}^{3} x^2\left(-\frac{1}{6}x+\frac{1}{2}\right)dx=\cdots\cdots=\frac{7}{6}$$

したがって

$$\{\sigma(X)\}^2=E(X^2)-\{E(X)\}^2=\frac{7}{6}-\left(\frac{2}{3}\right)^2=\frac{13}{18}$$

(3)　２つの解をα，βとしますと

$$\alpha+\beta=4,\ \alpha\beta=2X+1$$

また判別式 $D=3-2X$ ですから，２つの解がともに正であるための条件は

$$2X+1>0 \text{ かつ } 3-2X\geqq 0$$

です。これを解いて

$$-\frac{1}{2}<X\leqq\frac{3}{2}$$

ゆえに求める確率は

$$P\left(-\frac{1}{2}<X\leqq\frac{3}{2}\right)=\int_{-\frac{1}{2}}^{\frac{3}{2}} f(x)dx$$

$$=\int_{-\frac{1}{2}}^{0}\left(\frac{x}{2}+\frac{1}{2}\right)dx+\int_{0}^{\frac{3}{2}}\left(-\frac{x}{6}+\frac{1}{2}\right)dx$$

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

となります。では，もうひとつ。

■練習 2. 2つの定数 a，bは $a>0$, $0<b<1$ をみたし，確率変数Xの確率密度関数 $f(x)$ が

$$f(x)=\begin{cases} ax & (0\leqq x\leqq b) \\ \dfrac{ab}{b-1}x-\dfrac{ab}{b-1} & (b<x\leqq 1) \\ 0 & (x<0,\ x>1) \end{cases}$$

で与えられているとき，

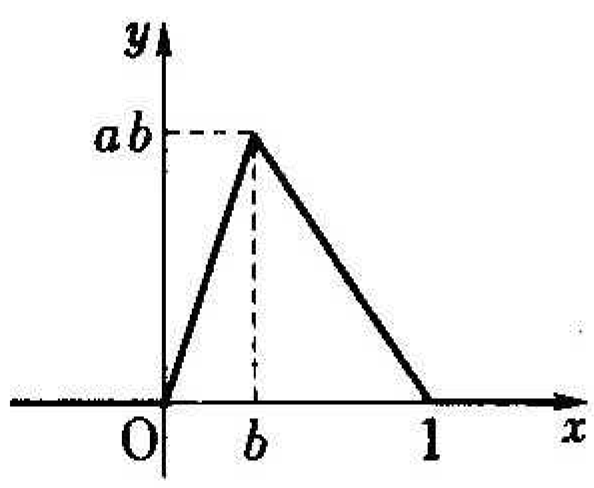
(1) abの値を求めよ。

(2) Xの平均値mを bで表せ。

(3) $b=\dfrac{1}{3}$ のとき，直線 $y=2x+X-\dfrac{1}{3}$ が円 $x^2+y^2=\dfrac{1}{45}$ に出合う確率を求めよ。

（高知大）

ヒント (1) まず確率密度関数のグラフをかいてみると右の図のようになります。これをx軸との囲む面積は1ですから，積分するまでもなし。



$$\frac{1}{2}b\cdot ab+\frac{1}{2}(1-b)\cdot ab=1$$

$$\therefore\ \frac{1}{2}ab=1$$

$$\therefore\ ab=2 \qquad \cdots\cdots\text{答}$$

(2) 平均の定義から

$$m=\int_0^b x(ax)dx+\int_b^1 x\left(\frac{ab}{b-1}x-\frac{ab}{b-1}\right)dx$$

$$=\frac{ab(b+1)}{6}=\frac{b+1}{3} \qquad \cdots\cdots\text{答}$$

(3) $b=\dfrac{1}{3}$ のとき $ab=2$ より

$$a=6$$

したがって

$$y=2x+X-\frac{1}{3} \text{ と } x^2+y^2=\frac{1}{45}$$

が共有点をもつためには

$$0\leqq X\leqq\frac{2}{3}$$

そこで求める確率は

$$\int_0^{\frac{2}{3}}f(x)dx=\int_0^{\frac{1}{3}}(6x)dx+\int_{\frac{1}{3}}^{\frac{2}{3}}(-3x+3)dx$$

$$=\frac{5}{6}$$

となります。どうです。もうひとつがんばりませんか。

■練習 3. 3点A$(a,0)$，C$(0,1)$，B$(b,0)$をこの順に結んだ2つの線分の作る折れ線がある。閉区間 $[a,\ b]$ のすべての値をとる確率変数 X の確率密度関数を $f(x)$ とする。ただし，$y=f(x)$ のグラフは上の折れ線で，$a<0$, $b>0$ とする。

(1) a，bの間にどんな関係があるか。

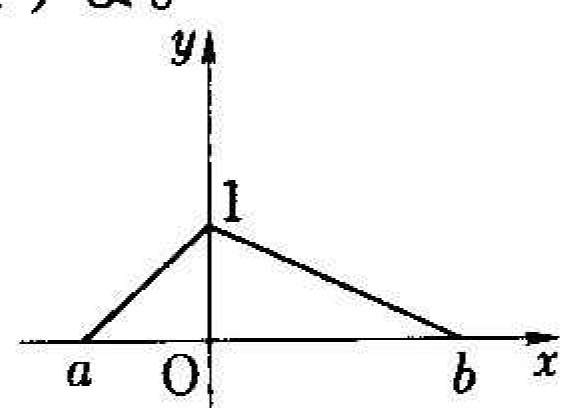
(2) 平均値 $E(X)$ を a の式で表せ。

(3) 分散 $V(X)$ を a の式で表せ。

(4) $\{E(X)-a\}^2=4V(X)$ となるようにaの値を定めよ。　（山梨大）

ヒント (1) $a<0$ という条件を忘れると，とんでもないことになりますよ。

さて，確率密度関数のグラフは右のようになりますから，この三角形の面積は1です。



$$\therefore\ \frac{1}{2}(-a)\cdot 1+\frac{1}{2}\cdot b\cdot 1=1$$

$$\therefore\ b-a=2$$

もちろん，$b>0$ ですから

$$a+2=b>0$$

$$\therefore\ a>-2$$

これと $a<0$ より

$$-2<a<0$$

あとは練習2と同様にやってみればよいのです。答は次の通り。

答 (1) $b-a=2$, $-2<a<0$

(2) $\dfrac{2}{3}(a+1)$ 　(3) $\dfrac{1}{18}(a^2+2a+4)$

(4) $a=-4+2\sqrt{3}$

＊　＊　＊

◎放物線分布では

1回目　年　月　日
2回目　年　月　日
3回目　年　月　日

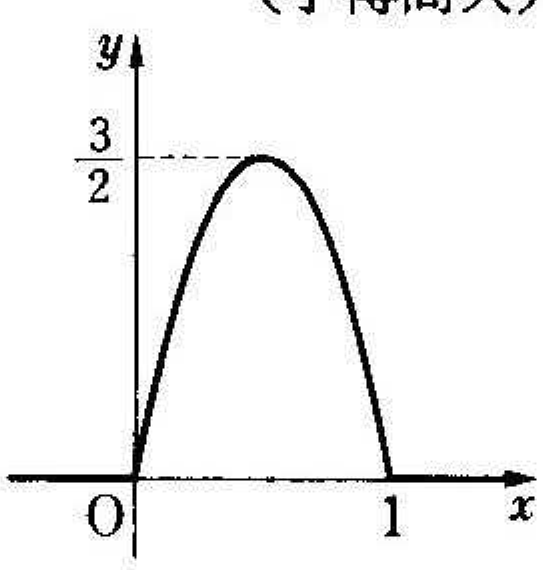
◆数学では直線の次が円，円の次は放物線ときまっているらしい。確率密度関数のグラフが放物線の場合はどうなるだろうか。

◆　ここでは，確率密度関数のグラフが，放物線である場合を考えてみることにしましょう。では，まず，これから：——

■練習 1. 確率Xの確率密度関数 $f(x)$ が

$$f(x)=\begin{cases}6(x-x^2) & (0\leqq x\leqq 1)\\ 0 & (x<0,\ 1<x)\end{cases}$$

で与えられるとき，Xの期待値は$\boxed{}$で，分散は$\boxed{}$である。　（小樽商大）



ヒント　定義に従って，そのとおりに計算すればよいのです。

まず，

$$E(X)=\int_0^1 x\cdot 6(x-x^2)dx$$

$$=6\int_0^1(x^2-x^3)dx$$

$$=6\left[\frac{x^3}{3}-\frac{x^4}{4}\right]_0^1=6\left(\frac{1}{3}-\frac{1}{4}\right)$$

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

次に：——

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

$$=\int_0^1 x^2\cdot 6(x-x^2)dx-\left(\frac{1}{2}\right)^2$$

$$=6\int_0^1(x^3-x^4)dx-\frac{1}{4}$$

$$=6\left[\frac{x^4}{4}-\frac{x^5}{5}\right]_0^1-\frac{1}{4}$$

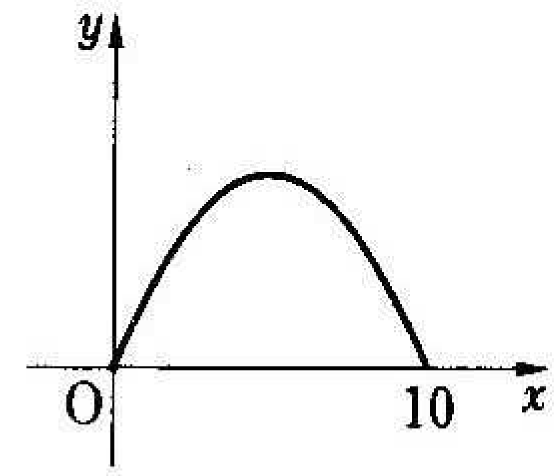
$$=\frac{3}{10}-\frac{1}{4}=\frac{1}{20}$$

となりましょう。

では，もうひとつ：——

■練習 2. 確率変数Xが区間［0，10］の任意の値をとることができ，その確率密度関数が，$f(x)=kx(10-x)$（kは定数）で与えられているとする。このとき，kの値を求めよ。また，確率 $P(3\leqq X\leqq 7)$ を求めよ。さらに，平均値と標準偏差を求めよ。

（旭川医大）



ヒント　$f(x)$ が確率密度関数ですから

$$\int_0^{10}kx(10-x)dx$$

$$=k\int_0^{10}(10x-x^2)dx$$

$$=k\left[5x^2-\frac{x^3}{3}\right]_0^{10}=k\frac{500}{3}$$

これが，1に等しくならなければなりません。

$$\therefore\quad \frac{500}{3}k=1,\qquad k=0.006$$

次に，

$$P(3\leqq X\leqq 7)=\int_3^7 0.006x(10-x)dx$$

$$=0.006\left[5x^2-\frac{x^3}{3}\right]_3^7=0.568$$

また，平均は定義によって

$$E(X)=\int_0^{10}0.006x^2(10-x)dx$$

$$=0.006\left[\frac{10}{3}x^3-\frac{x^4}{4}\right]_0^{10}$$

$$=0.006\times\frac{10000}{12}=5$$

最後に分散は，

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

$$=\int_0^{10}0.006x^3(10-x)dx-5^2$$

$$=0.006\left[\frac{10x^4}{4}-\frac{x^5}{5}\right]_0^{10}-5^2$$

$$=0.006\times\frac{100000}{20}-5^2=5$$

したがって，標準偏差は

$$D(X)=\sqrt{5}$$

となります。

■練習 3. 確率変数Xの確率密度関数 $f(x)$ がaを定数として

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

で与えられているとき

(1) aの値を求めよ。

(2) 確率 $P\left(X\leqq\frac{1}{4}\right)$ を求めよ。

(3) Xの平均 $E(X)$ と分散 $V(X)$ を求めよ。（福島県医大）

（解） (1) $f(x)$ は確率密度関数であるから

$$\int_0^1 f(x)dx=1$$

$$\therefore\quad \int_0^1 ax(1-x)dx=1$$

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

$$\therefore\quad a=6$$

(2) $$P\left(X\leqq\frac{1}{4}\right)=\int_0^{\frac{1}{4}}6x(1-x)dx$$

$$=6\left[\frac{x^2}{2}-\frac{x^3}{3}\right]_0^{\frac{1}{4}}=\frac{5}{32}$$

(3) $$E(X)=\int_0^1 xf(x)dx=\int_0^1 6x^2(1-x)dx$$

$$=6\left[\frac{x^3}{3}-\frac{x^4}{4}\right]_0^1=6\cdot\frac{1}{12}=\frac{1}{2}$$

また，

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

$$=\int_0^1 x^2f(x)dx-\left(\frac{1}{2}\right)^2$$

$$=6\int_0^1(x^3-x^4)dx-\frac{1}{4}$$

$$=6\cdot\frac{1}{20}-\frac{1}{4}=\frac{1}{20}$$

答 (1) 6 (2) $\frac{5}{32}$ (3) $\frac{1}{2}$，$\frac{1}{20}$

*　*　*

◆ ちょっと形が変わってこんなのもあります。──。これは $x=0$ と $x=0.5$ のところで不連続になっていますが気にすることはない。

■練習 4. Xの確率密度関数 $f(x)$ が

$$f(x)=\begin{cases}Ax^2+1 & \left(0\leqq x\leqq\frac{1}{2}\right)\\ 0 & \left(x<0,\ \frac{1}{2}<x\right)\end{cases}$$

で与えられているとき，$P(X\leqq a)=0.232$ となるaの値を定めよ。ただし，Aは正の定数とする。

（ヒント） $f(x)$ は確率密度関数ですから

$$\int_0^{\frac{1}{2}}f(x)dx=1$$

$$\therefore\quad \int_0^{\frac{1}{2}}(Ax^2+1)dx=\left[\frac{Ax^3}{3}+x\right]_0^{\frac{1}{2}}$$

$$=\frac{A}{24}+\frac{1}{2}=1$$

$$\therefore\quad \frac{A}{24}=\frac{1}{2}\qquad\therefore\quad A=12$$

そして，

$$P(X\leqq a)=\int_0^a(12x^2+1)dx$$

$$=\left[4x^3+x\right]_0^a=4a^3+a$$

$$=0.232$$

ですから，あとはこの3次方程式を解けばいいのでしょう。

シカシ，コレヲ解クノハ，イササカ抵抗ヲ感ズルネ。シカシ，常識的ニ考エルト，因数分解デトケルハズ。シカモ，0ヨリ大，0.5ヨリ小トナルト 0.2 ト 0.4 トシカアリエナイ。

割ってみると

$$(a-0.2)(4a^2+0.8a+1.16)=0$$

$$\therefore\quad a=0.2\qquad\cdots\cdots 答$$

あるいは次のように考えてもいいのです。

$$4a^3+a-0.232=0$$

両辺に 1000 をかけて

$$4000a^3+1000a-232=0$$

$10a=b$ とおくと

$$4b^3+100b-232=0$$

$$b^3+25b-58=0$$

これなら58の約数を入れてみて因数定理にもってゆく!!

1回目　年　月　日
2回目　年　月　日
3回目　年　月　日

◐ 三角関数や指数関数の入った分布では

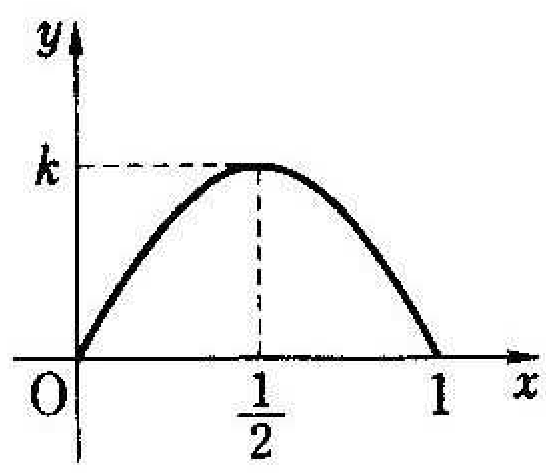
◆二項分布・正規分布・放物線分布といろいろやってきたが，ここでは三角関数にも遭遇する，というわけ。

◆ では，さっそくながらこれを：――

練習 1. 確率変数Xは区間［0, 1］の任意の値をとることができ，その確率密度関数は $k\sin\pi x$ であるという。このとき，kの値および $P(0\leqq X\leqq 0.5)$ を求めよ。

（滋賀医大）

ヒント　確率密度関数の性質から



$$\int_0^1 k\sin\pi x dx=1$$

でなければなりません。

ところが，左辺は

$$=\left[-\frac{k}{\pi}\cos\pi x\right]_0^1=\frac{2k}{\pi}$$

ですから

$$\frac{2k}{\pi}=1,\quad \therefore\quad k=\frac{\pi}{2}\quad \cdots\cdots 答$$

$$P(0\leqq X\leqq 0.5)=\int_0^{0.5}\frac{\pi}{2}\sin\pi x dx$$

$$=\frac{\pi}{2}\left[-\frac{\cos\pi x}{\pi}\right]_0^{0.5}=\frac{1}{2}\quad \cdots\cdots 答$$

練習 2. 確率変数Xの確率密度関数 $f(x)$ が $f(x)=\dfrac{a^2}{x^2+a^2}$ $(0\leqq x\leqq a)$ で与えられているとき，

(1) 定数aの値を求めよ。

(2) Xの平均値と分散を求めよ。

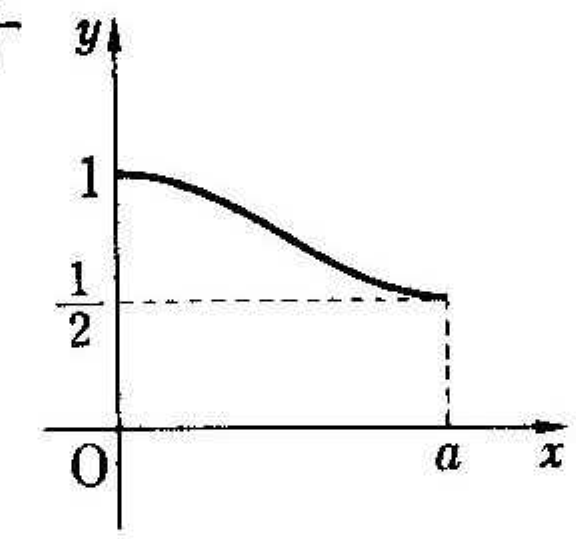
（福岡大）

ヒント　(1) $I=\displaystyle\int_0^a \frac{a^2}{x^2+a^2}dx$

において，$x=a\tan\theta$ $\left(0\leqq\theta<\dfrac{\pi}{2}\right)$ とおきなおすと，

$$I=\int_0^{\frac{\pi}{4}}\frac{1}{1+\tan^2\theta}\cdot\frac{a}{\cos^2\theta}d\theta=\int_0^{\frac{\pi}{4}}a d\theta=\frac{\pi}{4}a$$

これが1に等しいのですから



$$a=\frac{4}{\pi}$$

(2) 平均値mは

$$m=\int_0^a\frac{a^2x}{x^2+a^2}dx$$

$$=\left[\frac{a^2}{2}\cdot\log(x^2+a^2)\right]_0^a$$

$$=\frac{a^2}{2}\log 2$$

ところが

$$a=\frac{4}{\pi}$$

ですから

$$m=\frac{8}{\pi^2}\log 2$$

となります。次に分散です。

$$\sigma^2=\int_0^a\frac{a^2x^2}{x^2+a^2}dx-m^2$$

$$=\int_0^a\left(a^2-\frac{a^4}{x^2+a^2}\right)dx-m^2$$

$$=a^3-a^2-m^2$$

$$=\frac{16}{\pi^4}\{4\pi-\pi^2-4(\log 2)^2\}$$

となります。

*　　*　　*

◆ では，やや，総合的な問題にいきましょう。これは大阪府大の問題です。では：――

練習 3. ある電話局管内の電話の1回の通話時間（単位を分とする）は，確率変数Xで表され，その確率密度関数 $f(x)$ は

$$f(x)=\begin{cases} ae^{-\frac{x}{8}} & (0\leqq x<180)\\ 0 & (x\geqq 180)\end{cases}$$

である。ただし，aは定数である。一方通

話料は

$3(n-1)\leqq x<3n \quad (n=1,\ 2,\ \cdots\cdots)$

の通話時間に対して$10n$円である。このとき，次の問いに答えよ。

(1) 定数aの値を求めよ。

(2) 1回の通話時間の平均値を定積分で表し，その値（単位を分とする）を求めよ。

(3) 1回の通話料が$10n$円（$n=1,\ 2,\ \cdots\cdots,\ 60$）である確率を定積分で表し，その値を求めよ。

(4) 1回の通話料の平均値を級数で表し，その値（単位を円とする）を求めよ。

（大阪府大）

ヒント (1) 確率密度関数の性質から

$$\int_0^{180} ae^{-\frac{x}{3}}dx=a\left[-3e^{-\frac{x}{3}}\right]_0^{180}=1$$

$$\therefore\quad 3a(1-e^{-60})=1$$

$$\therefore\quad a=\frac{1}{3(1-e^{-60})} \qquad \cdots\cdots答$$

(2) 平均値 $m=\int_0^{180} xae^{-\frac{x}{3}}dx \qquad \cdots\cdots答$

$$=a\left[-3xe^{-\frac{x}{3}}\right]_0^{180}+3\int_0^{180} ae^{-\frac{x}{3}}dx$$

$$=\frac{1}{3(1-e^{-60})}\cdot\{-540e^{-60}\}+3$$

$$=\frac{3(e^{60}-61)}{e^{60}-1} \qquad \cdots\cdots答$$

(3) $P(3(n-1)\leqq x<3n)$

$$=\int_{3(n-1)}^{3n} ae^{-\frac{x}{3}}dx \qquad \cdots\cdots答$$

$$=a\left[-3e^{-\frac{x}{3}}\right]_{3(n-1)}^{3n}$$

$$=-3a(e^{-n}-e^{-n+1})$$

$$=\frac{(e-1)e^{-n}}{1-e^{-60}} \qquad \cdots\cdots答$$

(4) 求める平均値は

$$\sum_{n=1}^{60}10n\cdot\frac{(e-1)e^{-n}}{1-e^{-60}} \qquad \cdots\cdots答$$

$$=\frac{10(e-1)}{1-e^{-60}}\sum_{n=1}^{60}ne^{-n}$$

ところで，$\sum_{n=1}^{60}ne^{-n}$ はいわゆる循環級数といわれるもので，次のようにして求められます。

$$S=e^{-1}+2e^{-2}+3e^{-3}+\cdots+60e^{-60}$$

とおくと

$$e^{-1}S=\qquad e^{-2}+2e^{-2}+\cdots+59e^{-60}+60e^{-61}$$

となりますから，辺々相減じて

$$(1-e^{-1})S=e^{-1}+e^{-2}+\cdots+e^{-60}-60e^{-61}$$

$$=\frac{e^{-1}(1-e^{-60})}{1-e^{-1}}-60e^{-61}$$

$$\therefore\quad S=\frac{e(1-e^{-60})}{(e-1)^2}-\frac{60e^{-60}}{e-1}$$

かくして，求める値は

$$10\left(\frac{e}{e-1}-\frac{60}{e^{60}-1}\right)\text{円} \qquad \cdots\cdots答$$

注 このように，正規分布や二項分布以外にいろいろなものがつくられるわけです。たとえば，こんなのもあります。

≪区間［0, 1］の値のみをとる確率変数Xがあり，任意のx（$0\leqq x\leqq 1$）に対して $X\leqq x$ となる確率 $P_r(X\leqq x)$ が常に

$$P_r(X\leqq x)=\int_0^x kt^n(1-t)^2dt$$

と与えられる。ただし，nは自然数である。

(1) kをnで表せ。

(2) 確率変数Xの平均 $E(X)$ と分散 $V(X)$ をそれぞれnで表せ。（札幌医大）≫

ヒント ところで，この種の問題にみた瞬間ギョツとする，あっ，これは難しいぞ，と直感するわけです。しかし，そこでおどろいてはいけません。

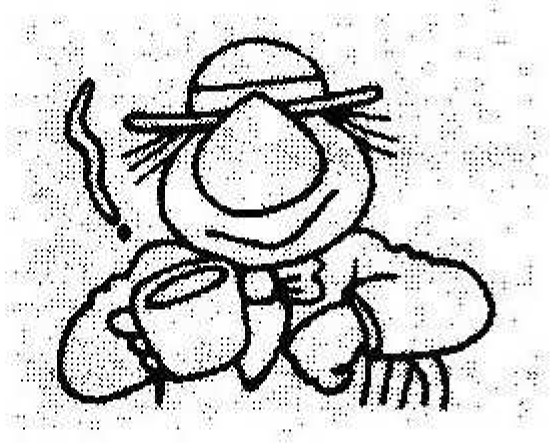
ではどうするか，まず，$n=1$ にしてやってみる。これならめんどうなし，できたら$n=2$としてやってみる，ついでだ，$n=3$としてやってみる，ここまでくると，誰もめんどうとは感じないものです。

あとは定義どおりやってみるといい。さあ，今スグやってみるべし。答を次にあげておきましょう。

答 (1) $k=\frac{1}{2}(n+1)(n+2)(n+3)$

(2) $E(X)=\frac{n+1}{n+4},\ V(X)=\frac{3(n+1)}{(n+4)^2(n+5)}$

* * *



確率分布で使う記号と蛇足

◆藤村の父親は神田橋のふきんで明治天皇に直訴した。それは，洋算の他に和算も教えるべきだ，というのであった。馴れの怖さ。

◆　一度覚えた記号や公式を変えることはかなり困難なものです。だから，いろいろな記号がハンランする。しかし，長い目でみると，自然に亡ぶものは亡んでゆくようです。ところで，確率分布のところにはいろんな記号が使われていて，それが入試問題にも反映しています。本書では無理に統一しませんでしたので，ここに一覧表をあげておきましょう。

名　称	記　号
確　率	P，p，P_r， $P(X \leqq x)$，$P_r(a \leqq X \leqq b)$
確率変数	X，Y，Z
平　均	m，$\bar{x}$，$E(X)$，μ
分　散	σ^2，s^2，$V(X)$，$\{\sigma(X)\}^2$
標準偏差	σ，s，$s.d.$，$D(X)$，$\sqrt{V(X)}$
標準化された変数	Z，U，t
標準正規分布の密度関数	$\phi(x)$，$\varphi(x)$
$\int_0^x \phi(x)\,dx$	J，$\Phi(x)$，$S(x)$

*　*　*

◆　また，分散から標準偏差を求めるときに平方根の求め方を知らないと困ることがあります。そこで右には，平方根の表をあげておくことにしましょう。なお，Nの平方根の近似値をuとするとき $\frac{1}{2}\left(u+\frac{N}{u}\right)$ はもっとよい近似値となります。

▶ 平方根の表 ◀

n	$\sqrt{n}$	n	$\sqrt{n}$
1	1.0000	51	7.1414
2	1.4142	52	7.2111
3	1.7321	53	7.2801
4	2.0000	54	7.3485
5	2.2361	55	7.4162
6	2.4495	56	7.4833
7	2.6458	57	7.5498
8	2.8284	58	7.6158
9	3.0000	59	7.6811
10	3.1623	60	7.7460
11	3.3166	61	7.8102
12	3.4641	62	7.8740
13	3.6056	63	7.9373
14	3.7417	64	8.0000
15	3.8730	65	8.0623
16	4.0000	66	8.1240
17	4.1231	67	8.1854
18	4.2426	68	8.2462
19	4.3589	69	8.3066
20	4.4721	70	8.3666
21	4.5826	71	8.4261
22	4.6904	72	8.4853
23	4.7958	73	8.5440
24	4.8990	74	8.6023
25	5.0000	75	8.6603
26	5.0990	76	8.7178
27	5.1962	77	8.7750
28	5.2915	78	8.8318
29	5.3852	79	8.8882
30	5.4772	80	8.9443
31	5.5678	81	9.0000
32	5.6569	82	9.0554
33	5.7446	83	9.1104
34	5.8310	84	9.1652
35	5.9161	85	9.2195
36	6.0000	86	9.2736
37	6.0828	87	9.3274
38	6.1644	88	9.3808
39	6.2450	89	9.4340
40	6.3246	90	9.4868
41	6.4031	91	9.5394
42	6.4807	92	9.5917
43	6.5574	93	9.6437
44	6.6332	94	9.6954
45	6.7082	95	9.7468
46	6.7823	96	9.7980
47	6.8557	97	9.8489
48	6.9282	98	9.8995
49	7.0000	99	9.9499
50	7.0711	100	10.0000