第15章　確率・統計

　ギャンブルは古来より人を夢中にさせるものであったようです．どっちに賭けるほうが得か，ギャンブラーは真剣に考えたことでしょう．3 次方程式の解の公式で名高い 16 世紀のイタリアの数学者カルダノ（Girolamo Cardano, 1501〜1576）は専門の賭博師でもあったとのことです．数学者や物理学者は，本人が賭博をやらなくても，ギャンブル好きな貴族に相談をもちかけられて確率の問題に関心をもったようです．地動説を唱えたガリレオ（Galileo Galilei, 1564〜1642，イタリア）は『さいころ賭博に関する考察』という論文を発表していますが，当時のイタリアではさいころ賭博が盛んで，ガリレオが受けた相談は“ 3 個のさいころを同時に振るとき，その目の数の和が 9 のときと 10 のときでは，それぞれ 6 通りで同じなのに，実際には 10 のほうが少し出やすいように思われるので調べてほしい ”というものでした（当時の計算法が誤りで，貴族の経験則のほうが正しいのです．後で議論しましょう）.

　“ 人間は考える葦(あし)である ”で有名な 17 世紀の数学者・哲学者パスカル（Blaise Pascal，1623〜1662，フランス）にもちかけられた相談はもっと生々しいようです．“ 2 人で賭けをしていて，その賭けは最終的に勝ったほうが全ての賭け金を貰えるとする．ところが（当局に踏み込まれて）この賭けを途中で止めたとき，途中の結果から判断すると賭け金をどのように配分すればよいか ”というものです（後で議論しましょう）．確率論に関する基礎研究はパスカルから始まったといわれており，彼は二項係数 ${}_nC_k$ に対して初めて用語「組合せ」を導入し，二項係数を三角形に配列した所謂(いわゆる)「パスカルの三角形」に理論的な基礎を与えました.

　確率論はその後多くの数学者の手を経て次第に内容が増し，フランスの数学者・物理学者ラプラス（Pierre Simon Laplace，1749〜1827）に至って，いわゆる「古典確率論」が 1812 年の大著『確率の解析的理論』に集大成されまし

た．彼が確率を定義するときに用いた用語“ 同程度に確からしい ”の根拠付けについては，その後の確率論において種々の議論を巻き起こしましたが，高校数学の範囲ではその直感的わかりやすさの故にあまり問題にしないで用いています．20 世紀になってから，確率論は，ロシアの数学者コルモゴロフ（Andrei Nikolaevich Kolmogorov，1903〜1987）によって，数学的に曖昧さがない公理論的な形で基礎付けがなされました．

現在，確率論は，ギャンブルの勝率はもちろん，人口統計・降水確率・分子の熱拡散運動（ブラウン運動）・各種保険の保険料の計算・年金の計算・遺伝学・商品の品質管理や販売計画・株価変動予測・その他の広い分野に応用されています．アメリカのディズニーランドの建設に当たっては，それを砂漠の真ん中に造って採算がとれるかどうかが大学の数学研究所に依頼され，あらゆる条件を詳細に検討した結果，確率的に採算がとれるとの結論を得て，GO サインが出ました．現在，企業が工場を建設したり，支店を出したり，新製品を開発したりする場合には確率計算を用いた検定が重要視されています．

§15.1　場合の数と確率

15.1.1　事象と確率

15.1.1.1　コイン投げ

1 個のコインを n 回投げて表が出る割合を考えましょう．コインはその円板という形状のゆえに，n 回投げるとその約半分は表，残りの半分は裏が出ますね（実際に 10 円玉を何度も投げてみて確かめましょう）．実際，投げる回数 n をいくらでも多くしていくと表が出る割合は $\frac{1}{2}$ に近づいていくことが予想され，実際そうなります．このとき，表が出る「経験的確率」または「統計的確率」は $\frac{1}{2}$ であるといいます．今の場合，表が出る確率を調べるためにコインを投げる試みを繰り返しましたね．一般に，ある現象の背後に潜む何らかの‘ 数学的な仕組みを探る目的で行う実験や観察 ’を **試行** といいます．また，今の場合コイン投げの試行の結果として表や裏が出ましたが，結果として起こる事柄を **事象** といいます．

さて，今の場合，表と裏が同時に出ることはありませんね．1 個のコイン投

げの試行では表が出る事象，裏が出る事象はそれぞれ個々に起こる事象なので‘それ以上細かく分けることができない事象’ですね．このような細分不可の事象を一般に **根元事象** といいます．異なる根元事象は同時に起こることはありません．一般に，同時に起こることがない 2 つの事象は **排反事象** と呼ばれます．また，表と裏のどちらが出てもよいという事象を考えるとその事象は必ず起こりますね．一般に，根元事象の全てを合わせた事象を **全事象** といい，その事象は必ず起こります．

ところで，試行の結果として個々に起こる根元事象は個々に観測される事柄でもあり，それを専門用語で「サンプル点」（標本点）といいます．よって，表や裏が出る根元事象はしばしばサンプル点として扱われます．

根元事象（サンプル点）の全体，つまり全事象は根元事象を要素とする集合 U と考えることができ，1 個のコイン投げの場合は $U = \{$表が出る，裏が出る$\}$ または簡略して $U = \{$表，裏$\}$ などのように表すことができます．この根元事象の集合 $U = \{$表，裏$\}$ をコイン投げ試行についての **サンプル空間**（標本空間）といいます．サンプル空間 $U = \{$表，裏$\}$ は 1 個のコインを投げるときに考えるべき全ての事柄を含んでいますね．一般の試行についてもサンプル空間，つまり全事象を考えることが確率の問題を考える基本になります．

さて，コイン投げの問題で表が出る確率が $\frac{1}{2}$ であることを理論的に考えましょう．そのためには‘コインには表と裏があるから半分は表が出る’ということに‘理屈をつければよい’わけです．そこで，全事象 $U = \{$表，裏$\}$ の各根元事象は‘同程度に起こりやすい’と仮定してみましょう．この仮定は自然なものとして受け入れられますね．その仮定を数値で表すには，根元事象の集合 $U = \{$表，裏$\}$ の要素の数 2，つまり根元事象の総数 2 を用いて，表が出るかまたは裏が出るかの各根元事象の起こる確率はそれぞれ $\frac{1}{2}$ であると仮定すればよいわけです．このような確率を「数学的確率」または「理論的確率」といいます．一般に，統計的確率と理論的確率が一致しそれらの間に矛盾がないことは後で示されます．確率の大きさを記号 P で表しましょう（P は probability（確率）の頭文字）．表が出る確率は $P = \frac{1}{2} = 0.5$ ですが，これは 1 回当たりの試行に換算すると表が 0.5 回出ることを表していますね．一般に，確率 P は‘1 回の試行に対して’ある事象が起こる割合を表します．

全事象 U の根元事象の総数が重要であることがわかったので，その数を一般に $n(U)$ で表しましょう．$U = \{表, 裏\}$ のとき $n(U) = 2$ です．このとき，表が出る確率を $P(表)$，裏が出る確率を $P(裏)$ とすると，各根元事象が同程度に起こりやすいという仮定は,

$$P(表) = P(裏) = \frac{1}{n(U)}$$

と表されます.

各根元事象の個数はいうまでもなく 1 ですが，事象を明示するためにその数を $n(表)$，$n(裏)$ などと書くと

$$P(表) = \frac{n(表)}{n(U)}, \qquad P(裏) = \frac{n(裏)}{n(U)}$$

と表されます．この表現を用いると，数学的確率は事象の数の比で表されることが明確になります.

根元事象の数を表すときに，$n(表) = 1$ を “表が出る **場合の数** は 1 通り”，$n(U) = 2$ を “全事象の場合の数は 2 通り” などという言い方もよくなされます．一般に，確率では異なるものが ‘ 何通りあるか ’ は重要で，事象 A に関する試行において，A の根元事象の数 $n(A)$ を「事象 A の場合の数」，試行の全事象 U の根元事象の総数 $n(U)$ を「全事象の場合の数」ということがあります．また，用語「場合の数」は，文字を並べるときなどに何通りの並べ方があるかなど，異なる種類の数を表すときにも用いられます.

次に，2 個のコインを同時に投げる試行の場合の数（根元事象の数）を調べましょう．もし同種のコインなら，問題を簡単にするために，色をつけて区別がつくようにしておきましょう．それらのコインを C_1，C_2 として，C_1 が表，C_2 が裏となる根元事象を (表, 裏) などと表しましょう．すると，サンプル空間，つまり全事象 U は，それぞれのコインの表・裏を考えると

$$U = \{(表, 表), (表, 裏), (裏, 表), (裏, 裏)\}$$

となりますね．根元事象 (表, 裏) と (裏, 表) は異なることに注意しましょう．

４つの根元事象の各々が 4 回に 1 回の割合で起こる，つまり $n(U)$ 回に 1 回の割合で起こることはいうまでもありません．ここで，片方のコインだけが表と

なる事象を A とすると

$$A = \{(表,\ 裏),\ (裏,\ 表)\}$$

です．よって，この試行においては各根元事象の起こり方は同程度に確からしいので，事象 A の場合の数 $n(A) = 2$，全事象の場合の数 $n(U) = 4$ より，4 回投げると 2 回の割合で事象 A が起こることがわかります．よって，事象 A の起こる数学的確率 $P(A)$ は $n(A)$ と $n(U)$ の比をとって

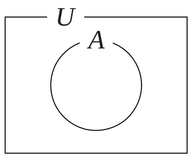
$$P(A) = \frac{n(A)}{n(U)} = \frac{2}{4} = \frac{1}{2}$$

となります．実際に 2 個の 10 円玉を 40 回投げてみると，(表, 表) と出たのが 9 回，(裏, 裏) が 10 回，(表, 裏) または (裏, 表) が 21 回となりました．この僅か 40 回の試行でも得られる事象 A の統計的確率は $\frac{21}{40} = 0.525$ となって数学的確率の結果にほぼ一致しました．

15.1.1.2　余事象・和事象・積事象

2 つの区別できるコイン C_1，C_2 を同時に投げる試行の全事象 U は

$$U = \{(表,\ 表),\ (表,\ 裏),\ (裏,\ 表),\ (裏,\ 裏)\}$$



このとき，片方のコインだけが表となる事象を A とすると

$$A = \{(表,\ 裏),\ (裏,\ 表)\}$$

でした．事象 A と全事象 U は根元事象の集合なので，それらはベン図を用いて表すことができます．

一般に，全事象 U のうち A でない事象を A の **余事象** といい，$\overline{A}$ で表します（$\overline{A}$ は A バーと読む）．ベン図でいうと，余事象 $\overline{A}$ は U から A をとり除いた部分です．今の場合，

$$\overline{A} = \{(表,\ 表),\ (裏,\ 裏)\}$$

となるので，$\overline{A}$ は同じ面が出る事象を表します．場合の数については明らかに

$$n(\overline{A}) = n(U) - n(A)$$

が成り立つので，どの根元事象も同程度に確からしいとき，事象 $\overline{A}$ が起こる確率は

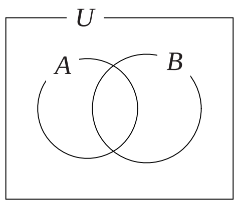
$$P(\overline{A}) = \frac{n(U) - n(A)}{n(U)} = 1 - P(A)$$

と表すことができます．

では，ここで問題です．区別がつく 5 個のコインを同時に投げたとき，少なくとも 1 個は表が出る確率 P を求めよ．ヒントはありません．

その事象を A としましょう．考慮すべき根元事象は (表, 表, 表, 表, 表) から (裏, 裏, 裏, 裏, 裏) まであり，よって全事象 U の場合の数 $n(U)$ は $2^5 = 32$ 通りありますね．そのうち事象 A が何通りあるか，それを直接数える気にはなりませんね．A の余事象 $\overline{A}$ は全て裏が出る場合の 1 通りしかないので，余事象の確率を用いて $P = 1 - \frac{1}{32} = \frac{31}{32}$ となります．もちろん，どの根元事象も同程度に確からしいと見なしています．

2 つのコインを投げる場合に戻りましょう．片方のコインだけが表になる事象



$$A = \{(表,\ 裏),\ (裏,\ 表)\}$$

に加えて，コイン C_1 が表となる事象 B を考えましょう：

$$B = \{(表,\ 表),\ (表,\ 裏)\}.$$

一般に，A または B が起こる事象を A と B の **和事象** といい，記号 $A \cup B$ で表します（$A \cup B$ は A または B と読みます）．今の場合，

$$A \cup B = \{(表,\ 表),\ (表,\ 裏),\ (裏,\ 表)\}$$

ですね．

また，A と B に共通する事象を A と B の **積事象** といい，記号 $A \cap B$ で表します（$A \cap B$ は A かつ B と読みます）．ベン図でいうと，積事象 $A \cap B$ は A と B の共通部分です．今の場合，$A \cap B$ は片方のコインだけが表でかつコイン C_1 が表となる事象ですから

$$A \cap B = \{(表,\ 裏)\}$$

です．

場合の数については，一般に，ベン図からわかるように

$$
\begin{aligned}
& n(A \cup B) = n(A) + n(B) - n(A \cap B) \\
\Leftrightarrow\ & n(A \cap B) = n(A) + n(B) - n(A \cup B)
\end{aligned}
$$

が成り立ちますね．したがって，確率については，どの根元事象も同程度に確からしいとすると

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

が成り立ちます．

なお，片方のコインだけが表になる事象 A に対して，B を両方のコインが共に表になる事象などとした場合には A と B は同時には起こらない排反事象になり，$A \cap B$ は要素（根元事象）がない空集合 $\emptyset$ になります．このような場合には $n(A \cap B) = 0$ となるので，和事象 $A \cup B$ の確率は

$$
P(A \cup B) = P(A) + P(B) \qquad （A \text{ と } B \text{ が排反事象のとき}）
$$

となります．

ここで問題です．区別がつく 3 個のコイン C_1，C_2，C_3 を同時に投げるとき，1 個だけ表が出る事象を A，コイン C_1 が表となる事象を B とする．このとき，A または B の事象が起こる確率を求めよ．ヒントなしです．

全事象 U の場合の数 $n(U)$ は $2^3 = 8$ 通りありますね．事象 A の場合の数 $n(A)$ はどのコインが表になるかで 3 通りあります．事象 B の場合の数 $n(B)$ は，コイン C_2，C_3 の表裏は関係がないので，$n(B) = 1 \cdot 2^2 = 4$ 通り．積事象 $A \cap B$ はコイン C_1 だけが表の場合になるから $n(A \cap B) = 1$ 通り．よって，

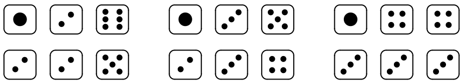
$$
P(A \cup B) = P(A) + P(B) - P(A \cap B) = \frac{3}{8} + \frac{4}{8} - \frac{1}{8} = \frac{3}{4}.
$$

15.1.1.3　ガリレオへのサイコロ相談

ガリレオが受けた相談“ 3 個のさいころを同時に振るとき，その目の数の和が 9 のときと 10 のときでは，それぞれ 6 通りで同じなのに，実際には 10 のほうが少し出やすいように思われるので調べてほしい ”に対する解答を考えま

しょう．もちろんサイコロは正しく作られていてどの目の出方も同じであるとします．

目の数の和が 9 になる事象を 和$_9$，10 になる場合を 和$_{10}$ としましょう．ガリレオの時代の人々は，事象 和$_9$ が起こる全ての場合をサイコロの目で表すと



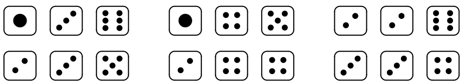
となるので，事象 和$_9$ が起こる場合の数は 6 通り（つまり，根元事象の数は 6）であるとして，それらはどれも同じ割合で起こると考えたわけです．この考えは 9 を 6 以下の 3 つの自然数の和に表すと 6 通りに表され：

$$9 = 1 + 2 + 6 = 1 + 3 + 5 = 1 + 4 + 4 = 2 + 2 + 5 = 2 + 3 + 4 = 3 + 3 + 3,$$

それらに同じ頻度を対応させるのと同じですね．同様に，事象 和$_{10}$ については

$$10 = 1 + 3 + 6 = 1 + 4 + 5 = 2 + 2 + 6 = 2 + 3 + 5 = 2 + 4 + 4 = 3 + 3 + 4$$

より，和$_9$ の場合と同じく



の 6 通りの目の出方があります．当時の貴族たちはそれら 12 通りの目の出方の各々は同じ割合で起こると考えたのでした．このように考えると事象 和$_9$ と 和$_{10}$ はほぼ同じ割合で起こりますね．

3 つのサイコロを 9400 回振る実験が実際になされ[1]，和$_9$ は 1072 回起こって全体の 11.4 %，和$_{10}$ は 1175 回で全体の 12.5 %でした．この実験からもわかるように，ガリレオに相談した貴族の経験通りに，和$_{10}$ のほうが起こりやすいことはほぼ疑いようがありません．当時の計算法はどこを誤ったのでしょうか，しばらくの間自ら考え，それから以下の議論に入っていきましょう．ヒント：先ほどのコインの問題ではコインを色分けしましたね．

3 個のさいころは区別がつかないように作られていますが，色をつけると結果は変わるでしょうか．変わりませんね．‘ サイコロの区別がつくようにして

[1] 文部省検定済み教科書『高等学校の確率・統計』（三省堂出版；1990 年発行）

振っても同じ結果を得る’ことは明らかでしょう．当時の計算法で見過ごしていたことは 3 つのサイコロを別物と考えることでした．

サイコロを区別することは場合の数を調べるときに決定的に重要です．例えば，先ほどの ⚀ ⚁ ⚅ と並べたサイコロはそれらの 3 個を区別してはいません．3 つのサイコロ (dice) を $\mathrm{D}_{\text{赤}}$，$\mathrm{D}_{\text{青}}$，$\mathrm{D}_{\text{黄}}$ などと区別して，振ったサイコロをそれらの順に並べ替えることにすると，⚀ ⚁ ⚅ に対して実際には

⚀ ⚁ ⚅　⚀ ⚅ ⚁　⚁ ⚀ ⚅　⚁ ⚅ ⚀　⚅ ⚀ ⚁　⚅ ⚁ ⚀

の 6 通りが現れます．つまり，当時の貴族のいう事象 ⚀ ⚁ ⚅ は実は場合の数を 6 通り（根元事象の数を 6）として考えるのが正しいのです．そのことはイメージ的には 6 通りに書かれた和

$$1+2+6,\quad 1+6+2,\quad 2+1+6,\quad 2+6+1,\quad 6+1+2,\quad 6+2+1$$

をある意味で別物と見なすことに当たるでしょうか．このようにこれまで 1 通りと見られていた場合の数を 6 通りに訂正する必要があります．このような訂正は出た目が全て異なる場合には全て同じです．練習問題として ⚀ ⚂ ⚄ の場合でそのことを確かめましょう．よって，出た目が全て異なる場合は 6 通りと数え直す必要があります．

次に，3 個のうち 2 個が同じ目の場合 ⚀ ⚃ ⚃ を $\mathrm{D}_{\text{赤}}$，$\mathrm{D}_{\text{青}}$，$\mathrm{D}_{\text{黄}}$ と区別して先ほどと同様に場合の数を数え直しましょう．実際には

⚀ ⚃ ⚃　⚃ ⚀ ⚃　⚃ ⚃ ⚀

と並べ直すのが正しく，今度は場合の数は 1 通りを 3 通りと訂正しなければなりませんね．⚁ ⚁ ⚄ でも同様です．

3 つとも同じ目が出た ⚂ ⚂ ⚂ の場合はどうでしょう．$\mathrm{D}_{\text{赤}}$，$\mathrm{D}_{\text{青}}$，$\mathrm{D}_{\text{黄}}$ の順に並べ替えてもやはり ⚂ ⚂ ⚂ と並びますね．よって，この場合の場合の数は 1 通りのままです．

以上のことから，事象 和_9 について場合の数を訂正したものを表にすると

⚀ ⚁ ⚅ : 6 通り，　⚀ ⚂ ⚄ : 6 通り，　⚀ ⚃ ⚃ : 3 通り，
⚁ ⚁ ⚄ : 3 通り，　⚁ ⚂ ⚃ : 6 通り，　⚂ ⚂ ⚂ : 1 通り

となるので，事象 和$_9$ の正しい場合の数 n(和$_9$) は

$$n(\text{和}_9) = 6 + 6 + 3 + 3 + 6 + 1 = 25\,\text{通り}$$

となりますね．

　同様に，事象 和$_{10}$ の場合に場合の数が n(和$_{10}$) = 27 通りとなることを確かめるのは練習問題です．ほとんど答ですが，ヒント：

⚀ ⚂ ⚅：6 通り，　⚀ ⚃ ⚄：6 通り，　⚁ ⚁ ⚅：3 通り，
⚁ ⚂ ⚄：6 通り，　⚁ ⚃ ⚃：3 通り，　⚂ ⚂ ⚃：3 通り．

　この問題では全事象 U の場合の数（根元事象の総数）$n(U)$ は，1 個のサイコロの目が 6 通りですから，D$_{赤}$，D$_{青}$，D$_{黄}$ の順に並べて考えるとわかるように

$$n(U) = 6 \times 6 \times 6 = 216\,\text{通り}$$

あります．それら 216 通りある根元事象のどれもが同程度に起こりやすいと考えられるので，各根元事象はサイコロを $n(U) = 216$ 回振ると 1 回の割合で起こります．したがって，出る回数はその事象の場合の数に比例し，n(和$_9$) = 25 通り，n(和$_{10}$) = 27 通りですから，事象 和$_{10}$ のほうが起こりやすいことがわかります．確率 P の表現を用いると

$$P(\text{和}_9) = \frac{n(\text{和}_9)}{n(U)} = \frac{25}{216} \fallingdotseq 0.116, \qquad P(\text{和}_{10}) = \frac{n(\text{和}_{10})}{n(U)} = \frac{27}{216} = 0.125$$

となります．これは先ほどの 9400 回サイコロを振って得られた実験結果，つまり 和$_9$ の場合が全体の 11.4 %，和$_{10}$ の場合が全体の 12.5 %に極めて近い値ですね．

　では，ここで問題です．3 個のサイコロを同時に振る試行を 1000 回繰り返す．そのうちサイコロの目が全て同じになるのは約何回か．ヒント：そんな場合は⚀ ⚀ ⚀～⚅ ⚅ ⚅です．その事象を A とすると $n(A) = 6$ 通りですから，1000 回のうち

$$P(A) \times 1000 = \frac{6}{216} \times 1000 \fallingdotseq 28\,\text{回}$$

しか出ませんね．

15.1.2　順列と組合せ

確率の本格的な計算に備えて場合の数の数え方を訓練しておきましょう．ある事象の場合の数の数え方が，あるものを並べるときの並べ方に同値であったり，あるものをとり出す場合のとり出し方に同値であることはよくあります．並べ方を「順列」，とり出し方を「組合せ」といいます．

15.1.2.1　重複（ちょうふく）順列

3 個のサイコロを 1 列に並べるとき，異なる並べ方が何通りあるか考えましょう．⚀⚁⚂ と ⚀⚂⚁ はもちろん違う並べ方です．実際に ⚀⚀⚀ から ⚅⚅⚅ まで並べてみるとわかるように，各サイコロには 6 通りの異なる目があり，それを 3 個並べる並べ方は $6^3 = 216$ 通りありますね．

n 種類の旗 $\boxed{1}$, $\boxed{2}$, $\cdots$, $\boxed{n}$ が何枚でもあります．この中から k 枚を選んで 1 列に並べ，旗の信号を作ります．このとき，同種の旗を繰り返して何回選んでもよいとします．旗の信号は何通り作れるでしょうか．実際に並べることを想像するとわかります．旗の各位置に立ってどの旗にしようかと考えると，旗は n 種類あるので，各位置で n 通りの旗の選び方がありますね．よって，旗を k 枚並べると n^k 通りの並べ方があります．

$$\underbrace{\overbrace{\boxed{h_1}\,\boxed{h_2}\cdots\boxed{h_k}}^{k\text{ 枚}}}_{\text{それぞれ } n \text{ 種類}}$$

一般に，異なる n 個のものの中から，同じものを何回でも選ぶことを認めて k 個とり出し，1 列に並べたものを，「n 個のものから k 個とる **重複順列**」といいます．その重複順列の総数を記号 ${}_n\Pi_k$ で表すと（Π は π の大文字）

$$ {}_n\Pi_k = n^k \qquad \text{（重複順列）}$$

です．

ではここで問題です．お年玉をたくさん貰った A 君は無駄使いせずに貯金しました．その後，A 君は新発売されたゲームソフトが欲しくなり貯金を下ろしに行ったのですが，4 桁の数字の暗証番号が思い出せません．しばらく考えて，千の位は 1 か 2，百の位は 0 でない偶数，十の位は 5 か 6，一の位は奇数であるところまでは思い出しました．忘れた数字は適当に入力するとして，彼は預金を下ろせるでしょうか．下ろせる確率を求めなさい．ただし，暗証番号

を三度続けて入力ミスするとそれ以上受け付けてもらえません．簡単のため，入力ミスをした後でまた同じ番号を入力するかもしれないとしましょう．ヒント：彼が思い出した部分は正しいとして，その中から無差別に数が選び出されます．

この問題の場合，数字を選ぶ全事象の場合の数は千の位で 2 通り，百の位で 4 通りなどと考えると，全部で $2 \times 4 \times 2 \times 5$ の 80 通りあります．その 80 通りの数のどれも平等に選ばれると考えます．入力する番号は 1 通りですが，三度のチャンスがあり，二度目・三度目に正しく入力する場合も考えます．よって，預金を下ろせる確率は

$$P = \frac{1}{80} + \frac{(80-1)\cdot 1}{80^2} + \frac{(80-1)^2\cdot 1}{80^3} \fallingdotseq \frac{3}{80} = 0.0375$$

です．この確率の値では預金を下ろすのは絶望的ですね．

15.1.2.2　順列

ア，イ，ウと書いたカード 3 枚 $\boxed{ア}\boxed{イ}\boxed{ウ}$ から 2 枚をとって並べる方法は，全部で

$$\boxed{ア}\boxed{イ}\quad\boxed{ア}\boxed{ウ}\quad\boxed{イ}\boxed{ア}\quad\boxed{イ}\boxed{ウ}\quad\boxed{ウ}\boxed{ア}\quad\boxed{ウ}\boxed{イ}$$

の 6 通りですね．

何故そうなったかを考えてみましょう．最初の文字を選ぶときは 3 通りありますが，次の文字を選ぶときは，各文字のカードが‘ 1 枚ずつしかない ’ために，$3-1=2$ 通りしか選べません．よって，並べ方の総数は

$$3 \times 2 = 6 \text{ 通り}$$

になります．これは，各カードが何枚でもある場合の結果 $3^2 = 9$ 通りとまったく違いますね．

n 人から k 人を選んで 1 列に並べることを考えましょう．最初の人を選ぶときは n 通り，2 人目は $n-1$ 通り，3 人目は $n-2$ 通り，$\cdots$，最後の k 人目の場合は $n-k+1$ 通りですから，並べ方の総数は

$$n(n-1)(n-2)\cdots(n-k+1) \text{ 通り}$$

となります．もし n 人全員を並べるとその総数は $n!$ 通りですね．

一般に，異なる n 個のものから，重複せずに異なる k 個のものをとり出して 1 列に並べたものを「n 個のものから k 個とる **順列**」といいます．その順列の総数を記号 ${}_n\mathrm{P}_k$ で表すと（P は permutation（順列）の頭文字です）

$$ {}_n\mathrm{P}_k = n(n-1)(n-2)\cdots(n-k+1) \qquad \text{（順列）} $$

ですね．

${}_n\mathrm{P}_k$ は階乗を用いて表すことができます．

$$ \begin{aligned} n(n-1)(n-2)\cdots(n-k+1) &= \frac{n(n-1)(n-2)\cdots(n-k+1)\cdot(n-k)!}{(n-k)!} \\ &= \frac{n!}{(n-k)!} \end{aligned} $$

より

$$ {}_n\mathrm{P}_k = \frac{n!}{(n-k)!} $$

となります．この形に表しておいて，§§14.8.3 で議論した解析的階乗関数に従って $0! = 1$,（負の整数）$! = \infty$ と約束すると，

$$ {}_n\mathrm{P}_0 = \frac{n!}{n!} = 1, \qquad {}_0\mathrm{P}_0 = \frac{0!}{0!} = 1, \qquad k > n \text{ のとき } {}_n\mathrm{P}_k = \frac{n!}{(n-k)!} = 0 $$

とすることになります．以下，そうしましょう．

では，ここで問題です．あるクラブの代表 5 人が大会に出場することになり，1 列に並んで記念写真を撮ることになった．代表 5 人の中には A 君と B さんがいて，A 君は B さんに密かに恋している．2 人が隣り合って並ぶ確率はいくらか．ただし，並ぶ位置は単なる偶然で決まるものとする．ヒント：一般に，確率の問題はさまざまな考え方ができ，解法は何通りもあるのが普通です．単にその問題が解けたと満足しないで，別の解法がないかと探るほうが後で役に立ってきます．一般的には場合分けが少ない解法が合理的ですが，場合分けが避けられない場合もあり，場合分けでミスをしない訓練も必要です．この問題では，A 君と B さんが隣り合う並び方の総数の求め方を工夫して，場合分けしないで済む解法を見つけましょう．

まず，全事象の場合の数（根元事象の総数）$n(U)$ は 5 人が 1 列に並ぶ順列の総数ですから，$n(U) = {}_5\mathrm{P}_5 = 120$ 通りありますね．このとき，根元事象は 5 人の並び方の各々であり，どの並び方も同程度に起こりやすいと考えます．よって，どの並び方も $\dfrac{1}{120}$ の確率で起こるとします．

次に，A 君と B さんが隣り合う並び方です．その場合の数を $n(AB)$ としましょう．2 人を A，B と略記し，隣り合って並んだ状態を AB や BA で表しましょう．すぐに気がつく解法は，並ぶ位置を 12345 とすると，AB または BA が 12，23，34，45 の位置にある場合で場合分けしておいて，残りの位置に他の 3 人が並ぶというものです：

AB345　1AB45　12AB5　123AB　および　AB を BA に変えたもの．

よって，2 人が並ぶ 4 通りの位置の各々に対して他の 3 人が残りの位置に並ぶ順列を考えればよいから，2 人が並ぶ場合の数は

$$n(AB) = 4 \times {}_3\mathrm{P}_3 \times 2 = 48 \text{ 通り}$$

です．この式で最後の $\times 2$ は 2 人の並び方が AB または BA の 2 通りあるためです．

この問題でより合理的な考え方は A，B の 2 人を 1 人と見なして，‘ 4 人 ’ が並ぶ順列を考えることです．すると，1 人と見なした 2 人を AB または BA に戻す手続きを最後に行って

$$n(AB) = {}_4\mathrm{P}_4 \times 2 = 48 \text{ 通り}$$

となります．

以上のことから，A 君と B さんが隣り合う確率は

$$P(AB) = \frac{n(AB)}{n(U)} = \frac{48}{120} = \frac{2}{5}$$

となります．A 君，期待してよいですよ．

ついでに，5 人がクラブの旗の周りに輪になって写真を撮る場合を考えましょう．この場合は 5 人の相対的位置のみを問題にします．5 人の並び方の総数を $n(U)$ として，輪をどこかで切って 5 人を 1 列に並ばせることを考えると，${}_5\mathrm{P}_5$ と $n(U)$ の間には

$${}_5\mathrm{P}_5 = n(U) \times 5$$

の関係が成り立つので

$$n(U) = \frac{{}_5\mathrm{P}_5}{5} = 4! \text{ 通り}$$

となりますね．

一般に，異なる n 個のものを円形に並ばせたものを **円順列** といい，その総数は

$$\frac{{}_n\mathrm{P}_n}{n} = (n-1)!$$

で与えられます．

なお，この場合に A 君と B さんが隣り合う確率は

$$P(AB) = \frac{3! \times 2}{4!} = \frac{1}{2}$$

となることを確かめましょう．ヒント：先の順列の問題で 2 人が 12 の位置にいる場合の並びを輪の並びに直すと考えるのがよいでしょう．

15.1.2.3　組合せ

abcd の 4 文字から 2 文字を単に選び出す方法の総数 N を考えましょう．実際に選び出すと

ab　ac　ad　bc　bd　cd

なので $N = 6$ 通りですね．これを論理的に考えてみましょう．abcd から 2 文字を選ぶことは，abcd から 2 文字とって並べる順列 ${}_4\mathrm{P}_2$ の中間段階に当たり，2 文字はまだ並べられていません．よって，N と ${}_4\mathrm{P}_2$ には

$${}_4\mathrm{P}_2 = N \times 2!$$

の関係があり，これから

$$N = \frac{{}_4\mathrm{P}_2}{2!} = \frac{4!}{2! \cdot 2!} = 6 \text{ 通り}$$

と求まります．

一般に，異なる n 個のものから，順序を考えずに k 個を選んで 1 組にしたものを「n 個のものから k 個をとる **組合せ**」といい，その総数を記号 ${}_n\mathrm{C}_k$ で表します（C は combination（組合せ）の頭文字です）．このとき，選んだ k 個を並べると ${}_n\mathrm{P}_k$ になるので

$${}_n\mathrm{P}_k = {}_n\mathrm{C}_k \times k!$$

が成り立ち，したがって

$${}_n\mathrm{C}_k = \frac{{}_n\mathrm{P}_k}{k!} = \frac{n!}{k! \cdot (n-k)!} \qquad \text{（組合せ）}$$

が得られます．この ${}_n\mathrm{C}_k$ は既に学んだ二項係数 ${}_n\mathrm{C}_k$ と同じものであり，それについては後で議論しましょう．

階乗の定義を拡張したので

$$ {}_n\mathrm{C}_0 = 1, \qquad {}_0\mathrm{C}_0 = 1, \qquad k > n \text{ のとき } {}_n\mathrm{C}_k = 0 $$

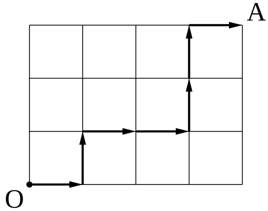
とすることになります．

ここで問題です．

$$ {}_n\mathrm{C}_k = \frac{n!}{k!(n-k)!} = \frac{n!}{(n-k)!(n-(n-k))!} = {}_n\mathrm{C}_{n-k}, \text{ よって } \quad {}_n\mathrm{C}_k = {}_n\mathrm{C}_{n-k} $$

が成り立ちますね．何故 ${}_n\mathrm{C}_k = {}_n\mathrm{C}_{n-k}$ なのでしょうか．納得がいく説明をせよ．ヒントは不要でしょう．

n 個の中から k 個を選ぶと，残される $n-k$ 個は決まりますね．よって，組合せは 1 つの集団を 2 つの集団に分ける方法を述べていると解釈できます．よって，n 個の集団を k 個の集団と $n-k$ 個の集団に分ける方法の総数が ${}_n\mathrm{C}_k$ で，$n-k$ 個の集団と k 個の集団に分ける総数が ${}_n\mathrm{C}_{n-k}$ ですからそれらは当然一致しますね．



右図の格子模様を道路と見立てて，地点 O から地点 A まで最短距離で行く方法の総数を考えましょう．図の矢印は 1 区画を進む方向を表し，O から A に最短で行く 1 つの例が示されています．まず，最短距離で行くための条件は何かと考えてみると，→ と進むのが 4 回，↑ が 3 回ですね．次に，通った路（みち）を区別するにはどうすればよいでしょうか．例で示したものは

1	2	3	4	5	6	7
→	↑	→	→	↑	↑	→

の順で進んでいます．矢印の上の数字が進む方向の順番を表しています．この表で → と ↑ の入れ替えを行うと通る路が異なってきますね．通る路を区別するのは順番を表す数字 1〜7 に対応する → と ↑ の違いというわけです．

以上の議論より，O から A に最短距離で行く方法の総数 N がわかります．数字 1〜7 から 4 つを選んで → を対応させ，残りの 3 つに ↑ を対応させます．

その総数は 7 個から 4 個をとる組合せの数なので

$$N = {}_7C_4 = 35 \text{ 通り}$$

の異なる行き方がありますね．もし歩き方に癖がない人が歩くとすると，どの最短路も $\dfrac{1}{35}$ の確率で選ばれますね．

応用問題です．文字 a が p 個，b が q 個，c が r 個の合計 n 個ある．それら n 個の文字を 1 列に並べる方法の総数 N を求めよ．ヒント：a を →，b を↑で置き換えてみると，前の最短距離の問題と考え方はほとんど同じでよいことがわかります．

文字を並べる位置を 1〜n として，1〜n から p 個を選んで文字 a を並べます．次に残りの $n-p$ 個の位置から q 個を選んで b を並べます．それで c を並べる位置も決まります．よって，このような並べ方の総数は組合せの数え方で求められ

$$N = {}_nC_p \cdot {}_{n-p}C_q = \frac{n!}{p!q!r!} \text{ 通り} \qquad (p+q+r=n)$$

となります．この問題は，‘同じものを並べる’場合には，順列の数え方ではなく，‘組合せの数え方’になることを示しています．

なお，この並べ方の問題は次の部屋割りの問題と答の数が一致します：n 人の人を A 室に p 人，B 室に q 人，C 室に残りの r 人を割り当てる方法の総数を求めよ．実際に解いてみて一致することを確かめましょう．

15.1.2.4　**重複組合せ**

お団子（だんご）6 個 ○○○○○○ を団子に目がない A，B，C の 3 兄弟で分ける方法の総数 N を考えましょう．団子を割ってはいけません．2 個ずつ分ければ平等なのですが，最悪の場合 1 個も貰えない人がいる場合もあるとします．しばらくの間考えてみましょう．団子は区別がつかないので，単純な組合せの問題ではありませんね．場合分けで解けないことはありませんが，勉強にはなりません．アイデアで勝負してみましょう．すんなり解けなかった人は一晩じっくり考えてみてください．この問題はその価値があります．

まず，例として A に 3 個，B に 2 個，C に 1 個分けた場合を図にしてみま

しょう.

$$\overbrace{○○○}^{\mathrm{A}}\ \overbrace{○○}^{\mathrm{B}}\ \overbrace{○}^{\mathrm{C}}$$

3 人のとり分はよくわかりますが，この図では計算ができるようにはなっていません．もう一工夫が要ります．団子を 1 列に並べたのがヒントになっています．3 人のとり分さえわかればよいのです.

今度は団子の間に間仕切りを入れてみましょう.

○○○▮○○▮○

これでも 3 人のとり分はわかりますね．間仕切りの位置を変えてみるとわかるように,‘ 間仕切りのおき方と分配の仕方が 1：1 に対応 ’ していますね．よって，間仕切りのおき方の総数が分配方法の総数 N になります.

このように団子と間仕切りを並べると計算ができる形になります．そのためにそれらの上に位置を示す数字を付けてみましょう．数字を ‘ 間仕切りの上にも付ける ’ のがミソです.

1	2	3	4	5	6	7	8
○	○	○	▮	○	○	▮	○

これで見えてきたでしょう．この問題は団子 6 個と間仕切り 2 個を 1 列に並べる方法の総数 N を求める問題になりました．それらをおく位置 1〜8 を考えて，その中から 2 ヵ所を選んで間仕切りに対応させると残りの位置は団子になります．よって，N は異なる 8 個から 2 個とる組合せの総数になりますね：

$$N = {}_8\mathrm{C}_2 = \frac{8 \cdot 7}{2!} = 28 \text{ 通り}.$$

解けてしまうと簡単ですが，団子に間仕切りという余計なものを加えて，かつそれらを同等に扱うという発想は誰にでもできるものではありません．確率は苦手だという人は ‘ 自分にはない考え方を吸収して自分のものにする ’ という気持ちで学ぶのがよいでしょう.

この分配の問題では誰かが 1 個も貰えない場合もあります．極端な話，A が 6 個とも独り占めする場合はこれです：○○○○○○▮▮.

では，ここで問題です．上の問題で，3 人とも少なくとも 1 個は貰えるとしたら，分配の方法は何通りあるか．ヒント：難しく考えないこと．

先に 1 個ずつ分けておいて，残りの 3 個を分ける方法を考えればよいわけです．間仕切りは 2 個だから，先の問題と同様にして

$$ {}_5C_2 = 10 \text{ 通り} $$

ありますね．

何百年も続くお菓子の福屋の饅頭（まんじゅう）松，竹，梅 は，19 代目がヤング向けに味つけを加味してからさらに評判がよくなり，売れ行きを伸ばしています．饅頭は一人 6 個限定ですが，松 竹 梅 の種類は問いません．饅頭の大好きな A さんは 6 個ずつ毎日違う組合せで買うことにしました．そのような買い方をすると買い終えるまでに何日かかるでしょうか．

団子の問題で練習したので，一発で片づけましょう．松 を 3 個，竹 を 2 個，梅 を 1 個買う例を

$$
\begin{array}{cccccccc}
1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\
\text{松} & \text{松} & \text{松} & \blacksquare & \text{竹} & \text{竹} & \blacksquare & \text{梅}
\end{array}
$$

と表せば納得できますね．団子 6 個を 3 人で分ける場合と完全に同じ数え方になります．よって，間仕切りをおく位置を毎日変えるとやり終えるのに

$$ {}_8C_2 = 28 \text{ 日}, $$

つまり，ほぼ 1 ヵ月かかりますね．

このタイプの問題を一般化しましょう．異なる n 個のものから，重複を許して k 個選んで 1 組にしたものを「n 個のものから k 個とる **重複組合せ**」といい，その総数を記号 ${}_nH_k$ で表します．${}_nH_k$ の値は下の例

$$
\begin{array}{ccccccccccccc}
1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & & & & & k+n-1 \\
1 & 1 & 1 & \blacksquare & 2 & 2 & \blacksquare & 3 & \cdots & \blacksquare & n & \cdots & n
\end{array}
$$

からわかるように k 個の◯と $n-1$ 個の間仕切りを並べるので，

$$ {}_nH_k = {}_{k+n-1}C_{n-1} = {}_{n+k-1}C_k $$

となります．

では，ここで問題です．生け花の先生が 5 種類の花を用いて花を生けるとき，種類に関係なく 6 本の枝を使いました．花の配置を考えないとすると，何通りの生け方が可能か．ノーヒントです．

花を 6 本並べ，花の種類を区別するのに 5 − 1 個の間仕切りをおくから

$$ {}_5\mathrm{H}_6 = {}_{5+6-1}\mathrm{C}_6 = {}_{10}\mathrm{C}_6 = {}_{10}\mathrm{C}_4 = 210 \text{ 通り} $$

の生け方がありますね.

15.1.2.5　二項定理とパスカルの三角形

数列の章で学んだ二項定理

$$ (x+y)^n = \sum_{k=0}^{n} {}_n\mathrm{C}_k x^{n-k} y^k $$

に現れる二項係数 ${}_n\mathrm{C}_k$ が組合せの考え方から導かれることを示しましょう．それには $(x+y)^n$ の展開がどのように行われるかを見れば済みます.

$$ (x+y)^n = \overbrace{(x+y)}^{1}\overbrace{(x+y)}^{2}\overbrace{(x+y)}^{3}\cdots\overbrace{(x+y)}^{n-1}\overbrace{(x+y)}^{n} $$

と展開前の因数 $(x+y)$ の位置に 1 から n まで番号を振っておきます．展開して得られる項は各因数 $(x+y)$ から x または y のどちらか一方のみを選び出して掛け合わせると得られますね．$x^{n-k}y^k$ 項は n 個の因数から k 個を選んで y をとり，残りの $n-k$ 個の因数から x を選んで掛け合わせると得られます．よって，$x^{n-k}y^k$ 項の係数は y をとる k 個の因数を選ぶ方法の総数で，その数はまさに n 個から k 個をとる組合せの数 ${}_n\mathrm{C}_k$ となりますね.

二項係数を求めるのに，かなり昔から用いられていた，いわゆる **パスカルの三角形** について議論しましょう．それは右図で表され，$(x+y)^n$ の展開項 $x^{n-k}y^k$ の係数を与えます．パスカルはこの三角形を組合せを用いて説明しました.

$$
\begin{array}{ccccccccc}
 & & & & 1 & & & & \\
 & & & / & & \backslash & & & \\
 & & & 1 & & 1 & & & \\
 & & / & \backslash & / & \backslash & & & \\
 & & 1 & & 2 & & 1 & & \\
 & / & \backslash & / & \backslash & / & \backslash & & \\
 & 1 & & 3 & & 3 & & 1 & \\
/ & \backslash & / & \backslash & / & \backslash & / & \backslash & \\
1 & & 4 & & 6 & & 4 & & 1
\end{array}
$$

1 段目の頂点 1 の位置から ／ や ＼ の路を通って最短距離で，例えば 3 の位置まで行くのに何通りの方法があるかを調べてみましょう．先に議論した

格子状道路の場合と同様に考えると，今度は ↙ や ↘ と進み，3 に行くには ${}_3\mathrm{C}_1 = {}_3\mathrm{C}_2 = 3$ 通りの方法がありますね．同様に 4 に行くには ${}_4\mathrm{C}_1 = {}_4\mathrm{C}_3 = 4$ 通り，6 に行くには ${}_4\mathrm{C}_2 = 6$ 通りの方法があります．

$$
\begin{array}{ccccccccc}
 & & & & {}_0\mathrm{C}_0 & & & & \\
 & & & {}_1\mathrm{C}_0 & & {}_1\mathrm{C}_1 & & & \\
 & & {}_2\mathrm{C}_0 & & {}_2\mathrm{C}_1 & & {}_2\mathrm{C}_2 & & \\
 & {}_3\mathrm{C}_0 & & {}_3\mathrm{C}_1 & & {}_3\mathrm{C}_2 & & {}_3\mathrm{C}_3 & \\
{}_4\mathrm{C}_0 & & {}_4\mathrm{C}_1 & & {}_4\mathrm{C}_2 & & {}_4\mathrm{C}_3 & & {}_4\mathrm{C}_4
\end{array}
$$

　このようにパスカルの三角形を読むと，それは一般に右図のように書くことができます．${}_0\mathrm{C}_0$ の位置から ${}_n\mathrm{C}_k$ まで行くのに ${}_n\mathrm{C}_k$ 通りの方法があるというわけです．

　さて，例えば ${}_4\mathrm{C}_2$ に最短で行くにはその直前の ${}_3\mathrm{C}_1$ か ${}_3\mathrm{C}_2$ を通らねばなりませんので，${}_4\mathrm{C}_2$ に行くには ${}_3\mathrm{C}_1 + {}_3\mathrm{C}_2$ 通りの方法があることになります．つまり，${}_4\mathrm{C}_2 = {}_3\mathrm{C}_1 + {}_3\mathrm{C}_2$ が成り立ちます．このことはそのまま一般化できて

$$
{}_n\mathrm{C}_k = {}_{n-1}\mathrm{C}_{k-1} + {}_{n-1}\mathrm{C}_k
$$

が成り立ちます．この公式は数列の章で二項係数を議論したときに既に得られていましたね．

　では，ここで問題です．二項定理を利用して公式：

$$
{}_n\mathrm{C}_0 + {}_n\mathrm{C}_1 + {}_n\mathrm{C}_2 + \cdots + {}_n\mathrm{C}_n = 2^n
$$

を導け．ヒントはないほうがよいでしょう．

$$
2^n = (1+1)^n = \sum_{k=0}^{n} {}_n\mathrm{C}_k
$$

で終りです．

　二項定理を利用して確率計算で役に立つ公式：

$$
\sum_{k=0}^{n} k\,{}_n\mathrm{C}_k x^{n-k} y^k = n(x+y)^{n-1} y
$$

を導きましょう．二項定理

$$
(x+y)^n = \sum_{k=0}^{n} {}_n\mathrm{C}_k x^{n-k} y^k
$$

の式は x, y についての恒等式なので，微分しても等号は成り立ちますね．よって，両辺を y で微分すると

$$n(x+y)^{n-1} = \sum_{k=0}^{n} k\,{}_n\mathrm{C}_k x^{n-k} y^{k-1}$$

が得られ，両辺に y を掛ければ終りです．得られた式はまた x, y についての恒等式になります．実際に用いられる公式は $y = p$，$x = 1 - p$ とおいて得られる

$$np = \sum_{k=0}^{n} k\,{}_n\mathrm{C}_k p^k (1-p)^{n-k}$$

です.

§15.2　確率

今まで学んできたことを吟味し，確率の基礎を体系化しましょう.

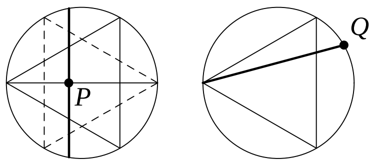
15.2.1　確率の基本性質

今まで学んできた確率ではある試行に対する全事象の根元事象の総数 $n(U)$ は有限であり，各根元事象は同程度に確からしく起こると考えました．しかしながら，例えば 100 人でジャンケンをして勝者が決まるまでの回数についての確率を考えると，いつまで経っても決まらない場合もあるので，$n(U) = \infty$ の場合も考えなければなりません．また，矢を的に当てる問題では当たる点の 1 つ 1 つが根元事象になるので，根元事象を連続量として扱う必要があります.

根元事象が同程度に起こりやすいという仮定は種々の問題をはらんでいます．鉛玉を埋め込んだ‘ いかさまサイコロ ’を振る場合には，出る目は同程度に確からしいとはできませんね．また，区別がつかない 2 つのコインを投げるとき，根元事象として 2 つとも表，1 つだけ表，2 つとも裏とした場合には同程度に起こりやすいの仮定は無効ですね.

その仮定を真に疑わせる問題が連続事象の場合に示されました：円とそれに内接する正三角形があるとき，円の任意の弦の長さが正三角形の 1 辺の長さよ

（ア）　（イ）



り大きくなる確率を考えます．このとき，弦の引き方によって答が変わります：（ア）正三角形の 1 つの辺に平行に弦を引く場合は，弦の中点 P は円の直径上にありますが，直径上のどの点にあるのも同程度に確からしいと考えると，確率は $\dfrac{1}{2}$ ですね．また，（イ）1 つの頂点を一方の端点とする弦を引く場合は，弦の他方の端点 Q は円周上にありますが，円周上のどの点にあるのも同程度に確からしいと考えると，確率は $\dfrac{1}{3}$ となります．つまり，‘ 何が根元事象であるか ’ または ‘ 何が同程度に確からしいか ’ が定かでないのです．この不定性は「ボーランドの逆説」として知られ，ラプラス流の確率論の見直しを促しました．

　やがて，これらの問題に対処できるように，確率論に公理的基礎付けがなされました．その理論で用いられる数学は積分を根本から見直すという高度なものですが，その出発点は以下のようなものです．ある試行を行ったときに起こる任意の事象 A を考えると，それは何も起こらない空事象 $\emptyset$ から必ず起こる全事象 U までさまざまなものが考えられます．よって，事象 A の起こる確率 $P(A)$ については

$$(1^\circ)\qquad P(\emptyset) = 0,\quad P(U) = 1$$

として

$$(2^\circ)\qquad \text{任意の事象 } A \text{ に対して}\quad 0 \leqq P(A) \leqq 1$$

が成り立つと仮定します（当たり前ですね）．

　また，根元事象はそれらのどれかが起こると他のものは起こらないので互いに排反事象になっています．そのことを一般化して，事象 A_1，A_2，A_3，$\cdots$，A_n（一般に $n = \infty$）が**互いに排反事象の**とき，それらの和事象の確率については，$A_k \cap A_l = \emptyset\ (k \neq l)$ なので

$$(3^\circ)\qquad P(A_1 \cup A_2 \cup A_3 \cup \cdots \cup A_n) = P(A_1) + P(A_2) + P(A_3) + \cdots + P(A_n)$$

が成り立つと仮定します．公理的確率論では基本的にはこれら 3 つの性質を確率の基本性質（公理）と考えます．

(3°) に現れる事象 A_k が根元事象 e_k で，n が有限なときは，$P(e_k)$ は根元事象の確率となります．このとき，必要があれば，同程度に確からしいという仮定をつけ加えて

$$P(e_k) = \frac{1}{n(U)}$$

とします．つまり，公理的確率論では‘同程度に確からしいという仮定は経験的法則’と見なします．

高校数学では，任意の事象 A，B に対して，(1°–3°) と集合の性質から導かれる余事象の確率の性質

(4°)　　$P(\overline{A}) = 1 - P(A)$

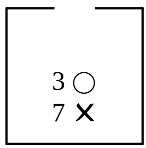
および，排反事象でない場合の和事象の確率の性質

(5°)　　$P(A \cup B) = P(A) + P(B) - P(A \cap B)$

を基本性質に加えます．

15.2.2　確率の積と条件付き確率

15.2.2.1　くじ引き



まず，くじ引きの問題に慣れておきましょう．箱の中に 10 本のくじが入っていて，そのうち 3 本は当たりくじ，7 本は外れくじです．くじを 2 本続けて引く試行を考えましょう．1 本目は当たる事象を $1_{\bigcirc}$，2 本目は当たる事象を $2_{\bigcirc}$ などと表しましょう．$1_{\bigcirc}$ は 2 本目の当たり・外れを問わない事象を意味し，$2_{\bigcirc}$ は 1 本目の当たり・外れはどうでもよい事象です．

始めに，事象 $1_{\bigcirc}$ が起こる確率を考えましょう．くじは 2 本引くのだから，まず 1 本目を引き，次に 2 本目を引くとしましょう．すると，この試行の全事象 U の場合の数 $n(U)$ は，10 本のくじは（裏に通し番号を打ったりして）区別がつくと考えたとき，10 本から 2 本を引く順列 ${}_{10}\mathrm{P}_2 = 10 \cdot 9$ となります．事象 $1_{\bigcirc}$ の場合の数 $n(1_{\bigcirc})$ は，1 本目が当たりで 2 本目はどちらでもよいから，$n(1_{\bigcirc}) = 3 \cdot 9$ 通りです．よって，その確率は，どのくじも同程度に引かれる

ので

$$P(1_{\bigcirc}) = \frac{3 \cdot 9}{10 \cdot 9} = \frac{3}{10}$$

となります．この結果は，たぶん君たちが予想していたように，くじを 2 本引いて 1 本目は当たる確率はくじを 1 本だけ引いたときに当たる確率 $\frac{3}{10}$ に一致しますね．

次に，1 本目の当たり外れを問わない事象 $2_{\bigcirc}$ の確率はどうでしょうか．場合の数 $n(2_{\bigcirc})$ は 1 本目の当たり外れを考慮して

$$n(2_{\bigcirc}) = 3 \cdot 2 + 7 \cdot 3 \text{ 通り}$$

となります．よって，その確率は

$$P(2_{\bigcirc}) = \frac{3 \cdot 2 + 7 \cdot 3}{10 \cdot 9} = \frac{3}{10}$$

となって $P(1_{\bigcirc})$ に一致しますね．これは‘くじ引きは引く順に関係なく公平である’ことの 1 つの証拠です．

くじ引きは公平であることの一般的な証明を行っておきましょう．それを場合分けの方法で行うのは無理なので，全事象 U の新たな設定を考えます．つまり，10 本のくじを順に全て引く試行を考え，くじは当たりと外れの区別だけをします．よって，その根元事象は，引く順を考慮すると，例えば

1	2	3	4	5	6	7	8	9	10
○	×	○	×	○	×	×	×	×	×

のように表すことができます．これは 1，3，5 本目が当たる場合の根元事象の例です．よって，全事象 U の場合の数は 1〜10 番から 3 ヵ所選んで当たりくじを対応させる組合せの総数 ${}_{10}\mathrm{C}_3$ 通りとなります：$n(U) = {}_{10}\mathrm{C}_3$．

k 本目は当たる場合の数 $n(k_{\bigcirc})$ は，k 番目に当たりくじを対応させておいて，残りの 9 ヵ所のうち 2 ヵ所に当たりくじを対応させればよいので，${}_9\mathrm{C}_2$ 通りあります：$n(k_{\bigcirc}) = {}_9\mathrm{C}_2$．したがって，$k$ 本目は当たる確率は

$$P(k_{\bigcirc}) = \frac{n(k_{\bigcirc})}{n(U)} = \frac{{}_9\mathrm{C}_2}{{}_{10}\mathrm{C}_3} = \frac{9 \cdot 8}{2!} \cdot \frac{3!}{10 \cdot 9 \cdot 8} = \frac{3}{10}$$

となり，この結果は k によりませんね．何番目に引いても当たる確率は同じです．

上の例で見たように，全事象の設定方法は問題の捉え方によって変わります．問題に応じてうまく設定すると解法が容易になりますね．順序を問わない場合は組合せの数え方ができるように設定するとよいでしょう．

15.2.2.2　確率の積・条件付き確率

さて，本題に入りましょう．3 本の当たりくじ・7 本の外れくじの入った箱から順に 2 本を引く試行で，1 本目も 2 本目も当たりくじを引く事象 $1_{\bigcirc} \cap 2_{\bigcirc}$ を調べましょう．くじは全て区別がつくとすると，その試行の全事象の場合の数は ${}_{10}\mathrm{P}_2$ 通り，事象 $1_{\bigcirc} \cap 2_{\bigcirc}$ の場合の数は $3 \cdot 2$ 通りありますから，その確率は

$$P(1_{\bigcirc} \cap 2_{\bigcirc}) = \frac{3 \cdot 2}{10 \cdot 9} = \frac{3}{10} \cdot \frac{2}{9} = P(1_{\bigcirc}) \cdot \frac{2}{9}$$

と表すことができます．上式の $\dfrac{2}{9}$ は，1 本目の当たりくじをとり去った残り 9 本のくじから，当たりくじを引く確率になっていますね．その確率については $\dfrac{2}{9} \fallingdotseq P(2_{\bigcirc})$ ですから，$\dfrac{2}{9}$ を $P_{1_{\bigcirc}}(2_{\bigcirc})$ と書くと，1 本目も 2 本目も当たりが出る確率は

$$P(1_{\bigcirc} \cap 2_{\bigcirc}) = P(1_{\bigcirc}) \cdot P_{1_{\bigcirc}}(2_{\bigcirc}) \quad \left(\Leftrightarrow \ \frac{3 \cdot 2}{10 \cdot 9} = \frac{3}{10} \cdot \frac{2}{9} \right)$$

と表されます．

上式は，一般に‘ 積事象の確率は確率の積の形で表すことができる ’ことを示唆しています．2 本とも当たる確率 $P(1_{\bigcirc} \cap 2_{\bigcirc})$ は 1 本目が当たる確率 $P(1_{\bigcirc})$ と‘ 1 本目が当たった状況の下で ’ 2 本目が当たる確率 $P_{1_{\bigcirc}}(2_{\bigcirc})$ との積であるということができます．このとき注意すべきことは，$P(1_{\bigcirc} \cap 2_{\bigcirc})$ に関する全事象がくじを 2 本引く試行についての事象であるのに対して，$P(1_{\bigcirc})$ と $P_{1_{\bigcirc}}(2_{\bigcirc})$ を求める際の実質的全事象は‘ くじを 1 本引く試行についての事象 ’と考えてよいことです．このために，積事象の確率を求める計算が著しく軽減されます．

同様に，2 本目は外れくじを引く事象を $2_{\times}$ と書くと，1 本目は当たりで 2 本目が外れとなる確率は

$$P(1_{\bigcirc} \cap 2_{\times}) = P(1_{\bigcirc}) P_{1_{\bigcirc}}(2_{\times})$$

と表すことができます．このとき $P_{1_{\bigcirc}}(2_{\times}) = \dfrac{7}{9}$ であることを示し，分数 $\dfrac{7}{9}$ の意味することを考えましょう（練習問題です）．

導出しましょう．事象 $1_{\bigcirc} \cap 2_{\times}$ の場合の数は $3 \cdot 7$ 通りあるから

$$P(1_{\bigcirc} \cap 2_{\times}) = \frac{3 \cdot 7}{10 \cdot 9} = \frac{3}{10} \cdot \frac{7}{9} = P(1_{\bigcirc})P_{1_{\bigcirc}}(2_{\times})$$

より明らかですね．この確率 $P_{1_{\bigcirc}}(2_{\times}) = \dfrac{7}{9}$ は当たりくじを 1 本抜いた 9 本のくじのうちから 7 本ある外れくじを引く確率です．よって，くじを 2 本続けて引いて 1 本目が当たりくじで 2 本目が外れである確率は，1 本目が当たりである確率 $P(1_{\bigcirc})$ と，1 本目が当たりである状況の下で，2 本目が外れる確率 $P_{1_{\bigcirc}}(2_{\times})$ との積で表されます．

以上の議論から，3 本目を引いたときに外れる事象を $3_{\times}$ とすると，事象 $1_{\bigcirc} \cap 2_{\times} \cap 3_{\times}$ が起こる確率は

$$P(1_{\bigcirc} \cap 2_{\times} \cap 3_{\times}) = \frac{3}{10} \cdot \frac{7}{9} \cdot \frac{6}{8} = \frac{7}{40}$$

のように積の形で求めることができます．これなら簡単ですね．

上の議論は直ちに一般化できます．2 つの事象 A，B の積事象 $A \cap B$ を考え，関係式

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

によって確率 $P_A(B)$ を定義します．$P_A(B)$ は，事象 A が起こったという条件の下で，事象 B が起こる **条件付き確率** と呼ばれます．その確率の値は一般には

$$P_A(B) = \frac{P(A \cap B)}{P(A)} \qquad \text{（条件付き確率）}$$

から求められます．このとき，事象 $A \cap B$ が起こる試行についての全事象 U の場合の数 $n(U)$ が定義でき，よって事象 A，$A \cap B$ の場合の数 $n(A)$，$n(A \cap B)$ が定義できる場合には

$$P_A(B) = \frac{P(A \cap B)}{P(A)} = \frac{n(A \cap B)}{n(U)} \cdot \frac{n(U)}{n(A)} = \frac{n(A \cap B)}{n(A)},$$

$$\text{よって} \quad P_A(B) = \frac{n(A \cap B)}{n(A)}$$

となるので，条件付き確率 $P_A(B)$ は，全事象を U から A に改めた場合の積事象 $A \cap B$ の確率という意味付けができます．先に議論した例題は全てこれに当てはまります．例えば，$P_{1_\bigcirc}(2_\times) = \dfrac{7}{9}$ では $n(1_\bigcirc) = 3 \cdot 9$，$n(1_\bigcirc \cap 2_\times) = 3 \cdot 7$ です．

では，ここで問題です．上のくじ引きの問題でくじを順に 3 本引くとき，始めの 2 本のうち 1 本だけ当たりくじを引き，3 本目は外れである確率を求めよ．ヒント：解法は何通りもありますが，場合分けをしなくて済む方法があります．当たりくじは 1 本目でも 2 本目でもよいので，始めの 2 本を 2 本一緒にまとめて引いても構いません．すると，始めの 2 本を引く試行についての全事象の場合の数は，くじは全て区別がつくとして，10 本から 2 本を引く組合せの総数 ${}_{10}\mathrm{C}_2$ と見なすことができます．

始めの 2 本を引くことの全事象の場合の数を組合せの考えで数えた場合には，そのうちの 1 本が当たる事象の場合の数も組合せの考えで求めます．1 本当たり・1 本外れなので，その場合の数は ${}_3\mathrm{C}_1 \cdot {}_7\mathrm{C}_1$ です．3 本目を引くときは当たりくじ・外れくじが 1 本ずつ減っていることを考慮すると，求める確率は，確率の積の性質を利用して

$$P = \frac{{}_3\mathrm{C}_1 \cdot {}_7\mathrm{C}_1}{{}_{10}\mathrm{C}_2} \cdot \frac{6}{8} = \frac{7}{20}$$

となります．この方法はぜひマスターしましょう．

この問題を 3 本のくじを順に引くという考え方で解いてみます．場合分けと確率の積の性質を完全に用いると

$$P = \frac{3}{10} \cdot \frac{7}{9} \cdot \frac{6}{8} + \frac{7}{10} \cdot \frac{3}{9} \cdot \frac{6}{8} = \frac{7}{20}$$

です．3 本目を引くときに確率の積の性質を用いると

$$P = \frac{3 \cdot 7 + 7 \cdot 3}{{}_{10}\mathrm{P}_2} \cdot \frac{6}{8} = \frac{7}{20}$$

となります．確率の積を用いなければ

$$P = \frac{3 \cdot 7 \cdot 6 + 7 \cdot 3 \cdot 6}{{}_{10}\mathrm{P}_3} = \frac{7}{20}$$

となります．

では，条件付き確率の問題をやってみましょう．上のくじ引きの問題で，既に当たりくじが 2 本引かれているとき，5 番目に引く人が当たりくじを引く確率を求めよ．ヒント：難しく考える必要はありません．

確率の積の性質を利用します．5 番目に引くのだから，残っているくじは 6 本，そのうち当たりくじは 1 本だから，求める確率は $P = \dfrac{1}{6}$ ですね．

最後に，条件付き確率問題の傑作といわれている古典的大学入試問題をじっくり考えてみましょう：5 回に 1 回の割合で帽子を忘れる癖のある K 君が，正月に A，B，C の 3 軒をこの順に年始回りをして家に帰ったとき，帽子を忘れてきたことに気がついた．2 軒目の家 B に忘れてきた確率を求めよ．ヒント：‘帽子を忘れてきた’という条件の下でその確率を求めます．この問題では条件付き確率の関係式 $P(A \cap B) = P(A)P_A(A \cap B)$ を用いないと解けません．

3 軒のうちのどこかで帽子を忘れてきた事象を X，特に A，B，C で忘れた事象をそれぞれ A，B，C で表しましょう．この問題の全事象はこの 3 つの事象に加えてどこにも忘れてこない事象 $\overline{\mathsf{X}}$ からなります．$\mathsf{X} = \mathsf{A} \cup \mathsf{B} \cup \mathsf{C}$，また $\mathsf{X} \cap \mathsf{B} = \mathsf{B}$ であることに注意しましょう．よって，条件付き確率の関係式から

$$P(\mathsf{X} \cap \mathsf{B}) = P(\mathsf{X})P_{\mathsf{X}}(\mathsf{X} \cap \mathsf{B}) \iff P(\mathsf{B}) = P(\mathsf{X})P_{\mathsf{X}}(\mathsf{B})$$

が得られます．上式の $P(\mathsf{B})$ は単に B で帽子を忘れてきた確率（その全事象は忘れてこない事象 $\overline{\mathsf{X}}$ を含む）で，忘れてきたという前提条件は考慮されていません．そのことを考慮した確率は $P_{\mathsf{X}}(\mathsf{B})$ のほうで，これが求めるものです．

この問題を場合の数で考えるのは難しいので，$P_{\mathsf{X}}(\mathsf{B})$ を直接求める代わりに $P(\mathsf{B})$ と $P(\mathsf{X})$ を利用します．$\mathsf{X} = \mathsf{A} \cup \mathsf{B} \cup \mathsf{C}$ において，事象 A，B，C は互いに排反事象なので

$$P(\mathsf{X}) = P(\mathsf{A}) + P(\mathsf{B}) + P(\mathsf{C})$$

が成り立ちます．K 君は 5 回に 1 回の割合で帽子を忘れる癖があり，A，B，C の順で回るので，確率の積の性質を用いると

$$P(\mathsf{A}) = \frac{1}{5}, \qquad P(\mathsf{B}) = \frac{4}{5} \cdot \frac{1}{5}, \qquad P(\mathsf{C}) = \frac{4}{5} \cdot \frac{4}{5} \cdot \frac{1}{5}$$

が得られます．$P(\mathsf{B})$ の $\dfrac{4}{5}$ は A で忘れなかったことを表します．$P(\mathsf{C})$ も同様．

よって，

$$P(\mathsf{X}) = \frac{1}{5} + \frac{4}{5} \cdot \frac{1}{5} + \frac{4}{5} \cdot \frac{4}{5} \cdot \frac{1}{5} = \frac{61}{125}$$

となるので

$$P_{\boldsymbol{\times}}(\mathrm{B}) = \frac{P(\mathrm{B})}{P(\boldsymbol{\times})} = \frac{4}{25} \cdot \frac{125}{61} = \frac{20}{61}$$

が得られます．

　なお，この条件付き確率を $P(\mathrm{B})$，$P(\boldsymbol{\times})$ を参考にして

$$P_{\boldsymbol{\times}}(\mathrm{B}) = \frac{4 \cdot 1 \cdot 5}{1 \cdot 5^2 + 4 \cdot 1 \cdot 5 + 4^2 \cdot 1}$$

と書くと場合の数を用いた考察ができます．目の数が 1〜5 しかない 5 面体のサイコロを考え，1〜4 の目が出たら忘れないとし，5 の目が出たら忘れるとします．K 君が立ち寄った家で悪魔がこのサイコロを振って，忘れる・忘れないを決めます．この試行の全事象は 5^3 通りあり，上式の分子が場合の数 $n(\mathrm{B})$，分母が $n(\boldsymbol{\times})$ です．ただし，上式に現れる 5，例えば分母の第 1 項 $1 \cdot 5^2$ の 5 については，A で既に忘れてしまったので，B，C ではどの目でもよいと考えます．分母の $n(\boldsymbol{\times})$ にはどこにも忘れてこない事象 $\overline{\boldsymbol{\times}}$ の場合の数 4^3 が含まれていませんね．このことが忘れてきたという条件を表しています．全事象の場合の数が $n(\boldsymbol{\times}) + 4^3 = 5^3$ であることを確かめましょう．

15.2.3　独立事象の確率

15.2.3.1　事象の独立

　1 つのサイコロを 2 回振る試行で，1 回目に ⚀ が出る事象を $1_{\text{⚀}}$，2 回目に ⚁ が出る事象を $2_{\text{⚁}}$ として，積事象 $1_{\text{⚀}} \cap 2_{\text{⚁}}$ が起こる確率を考えます．その確率を求めなさいといわれると，全員が

$$P(1_{\text{⚀}} \cap 2_{\text{⚁}}) = \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}$$

と解答するでしょう．これはもちろん正解です．以下，厳密な扱いをして上式の意味を考えましょう．

　サイコロを 2 回振る試行ですから，全事象の場合の数は 6^2 通りあります．このとき，事象 $1_{\text{⚀}}$ は，2 回目の目を指定しないので，$1 \cdot 6$ 通りあります．同様に，事象 $2_{\text{⚁}}$ の場合の数は $6 \cdot 1$ 通りあります．よって，それらの事象が起こる確率は

$$P(1_{\text{⚀}}) = \frac{1 \cdot 6}{6^2} = \frac{1}{6}, \qquad P(2_{\text{⚁}}) = \frac{6 \cdot 1}{6^2} = \frac{1}{6}$$

と，当然の結果になります．また，積事象 $1_{\text{⚀}} \cap 2_{\text{⚁}}$ の場合の数は 1 通りなので，その確率は

$$P(1_{\text{⚀}} \cap 2_{\text{⚁}}) = \frac{1}{6^2} = \frac{1}{6} \cdot \frac{1}{6}$$

となり，

$$P(1_{\text{⚀}} \cap 2_{\text{⚁}}) = P(1_{\text{⚀}}) \cdot P(2_{\text{⚁}})$$

の関係が成り立ちます．

一方，積事象の確率は条件付き確率を用いて積の形に書けて

$$P(1_{\text{⚀}} \cap 2_{\text{⚁}}) = P(1_{\text{⚀}}) \cdot P_{1_{\text{⚀}}}(2_{\text{⚁}})$$

と表されますね．これと先ほどの結果から

$$P_{1_{\text{⚀}}}(2_{\text{⚁}}) = P(2_{\text{⚁}})$$

が成り立ちます．この式の意味は，1 回目に ⚀ が出たときに 2 回目に ⚁ が出る確率は（1 回目に出た目を問わずに）2 回目に ⚁ が出る確率に等しいというわけですから，1 回目に出た目は 2 回目に出る目に影響しないということです．至極当然ですね．

このようにある事象が他の事象に影響を与えないことはよくあることです．先に議論したくじ引きの問題でも，もし引いたくじを元に戻して引き直すとすれば，前の結果は後の結果に影響しませんね．一般に，事象 A が事象 B に影響を及ぼさないことは

$$P_A(B) = P(B), \quad \text{または} \quad P(A \cap B) = P(A) \cdot P(B)$$

として表され，B は A に **独立** であるといいます．このとき，$A \cap B = B \cap A$ に注意して

$$P_A(B) = P(B) \iff \frac{P(A \cap B)}{P(A)} = P(B) \iff \frac{P(B \cap A)}{P(B)} = P(A) \iff P_B(A) = P(A),$$

$$\text{よって} \quad P_A(B) = P(B) \iff P_B(A) = P(A)$$

が導かれるので，B は A に独立なとき A は B に独立になります．したがって，

$$P(A \cap B) = P(A) \cdot P(B) \qquad \text{（}A \text{ と } B \text{ は独立）}$$

が成り立つとき, ‘ 事象 A と B は独立 ’ であると定義するほうがよいでしょう．事象 A と B が独立でないとき，A と B は **従属** であるといいます.

なお，先ほどのサイコロ振りですが，サイコロを 1 回振ることを 1 つの試行と見なし，続けて振ることを試行を繰り返すということがあります．この場合，前の試行の結果と後の試行の結果が互いに影響しないことをこれらの ‘ 試行は独立 ’ であるということもあります．一般に，2 つの試行を行うとき，一方の試行の結果が他方の試行の結果とは無関係に決まるとき，それらの試行は独立であるといいます.

2 つの試行 S（例えば，サイコロ振り）と T（例えば，コイン投げ）を考え，試行 S における任意の事象を A_{S}，試行 T における任意の事象を A_{T} としましょう．このとき，試行 S，T を合わせたものを 1 つの試行 U と考えることができ，A_{S}，A_{T} は試行 U の事象と見なすことができます．このとき，A_{S}，A_{T} を合わせ考えた事象は試行 U の積事象 $A_{\mathrm{S}} \cap A_{\mathrm{T}}$ として表されます．その確率は,一般に

$$P(A_{\mathrm{S}} \cap A_{\mathrm{T}}) = P(A_{\mathrm{S}}) \cdot P_{A_{\mathrm{S}}}(A_{\mathrm{T}})$$

となり，特に試行 S，T が独立ならば

$$P(A_{\mathrm{S}} \cap A_{\mathrm{T}}) = P(A_{\mathrm{S}}) \cdot P(A_{\mathrm{T}})$$

となります．なお，異なる試行を合わせて 1 つの試行と見なしたり，その逆に 1 つの試行を複数の事象に分解することは，特に断りなく行われます．その辺は状況から判断しましょう.

ここで練習問題です．A，B の 2 人がジャンケンをします．3 回行って全てアイコになる確率を求めよ．ヒント：各回のジャンケンを 1 つの試行と考えると，ジャンケンは独立な試行です.

グー・チョキ・パーの組合せを考えると（グー，グー）から（パー，パー）までの 9 通りで，そのうちアイコになるのは 3 通り．よって，アイコになる確率は $\dfrac{1}{3}$ です．ジャンケンは独立試行なので，3 回共アイコである確率は

$$P = \left(\frac{1}{3}\right)^3 = \frac{1}{27}$$

ですね.

15.2.3.2　反復試行の確率

コインやサイコロを投げる試行を繰り返す場合は同じ条件で繰り返しますね．このような繰り返しの試行を **反復試行** といい，各回の試行は独立な試行になります．

サイコロを n 回振る反復試行において，$\boxed{\bullet}$ が n 回のうち k 回出る確率 $P(k)$ を求めましょう．まず，始めの k 回で $\boxed{\bullet}$ が出てしまい，残りの回には出ない確率 P_0 は，サイコロ振りが独立試行なので

$$P_0 = \left(\frac{1}{6}\right)^k \left(\frac{5}{6}\right)^{n-k}$$

ですね．次に，1〜n 回のうちの任意に選ばれた k 回で $\boxed{\bullet}$ が出て残りの $n-k$ 回では出ない場合を考えましょう．例えば，$\boxed{\bullet}$ 以外の目を $\boxed{\overline{\bullet}}$ として

$$\begin{array}{ccccccc} 1 & 2 & 3 & \cdots & & \cdots & n \\ \boxed{\overline{\bullet}} & \underbrace{\boxed{\bullet}}_{1\text{ 個目}} & \boxed{\overline{\bullet}} & \cdots & \underbrace{\boxed{\bullet}}_{k\text{ 個目}} & \cdots & \boxed{\overline{\bullet}} \end{array}$$

です．この場合の確率も $\boxed{\bullet}$ の出る回に関係なく $\left(\frac{1}{6}\right)^k \left(\frac{5}{6}\right)^{n-k} = P_0$ となりますね．よって，n 回のうちから $\boxed{\bullet}$ が出る k 回を選ぶ方法の総数は ${}_n\mathrm{C}_k$ 通りあり，それら ${}_n\mathrm{C}_k$ 通りある事象は互いに排反事象だから，求める確率 $P(k)$ はそれらの確率の単なる和になります：

$$P(k) = {}_n\mathrm{C}_k \left(\frac{1}{6}\right)^k \left(\frac{5}{6}\right)^{n-k}.$$

この結果は直ちに一般化できます．反復試行において，1 回の試行で事象 A の起こる確率を p とします．この試行を n 回繰り返すとき，A が k 回起こる確率 $P(k)$ は

$$P(k) = {}_n\mathrm{C}_k\, p^k (1-p)^{n-k} \qquad (k = 0,\ 1,\ 2, \cdots,\ n)$$

で与えられます．

1	2	3	4	5
$\overline{\text{勝}}$	勝	勝	勝	$\overline{\text{勝}}$

では，ここで問題です．A，B の 2 人がジャンケンを 5 回した．(1) A が 4 回以上勝つ確率を求めよ．(2) A が少なくとも 1 回は勝つ確率を求めよ．(3) A が 3 回勝つ，しかも連続して勝つ確率を求めよ．(4) A が始

めの 3 回は連続して勝つ確率を求めよ．(5) A が 3 回以上連続して勝つ確率を求めよ．ヒント：1 回のジャンケンでどちらかが勝つ，負ける，引き分ける確率は全て $\frac{1}{3}$ です．(2) は利口に対処します．(4) は問題の意味をとり違えないように．(5) は場合分けで間違えないように．

(1) A が 4 回または 5 回勝つ場合ですね．勝たない場合は負けか引き分けです．よって，その確率は

$$P = {}_5C_4\left(\frac{1}{3}\right)^4\frac{2}{3} + {}_5C_5\left(\frac{1}{3}\right)^5 = \frac{11}{243}.$$

(2) 1 回勝つ〜5 回勝つ場合の確率を求めて加える気にはならないでしょう．「少なくとも 1 回は勝つ」の余事象は「1 回も勝てない」です．よって，余事象の確率を用いて

$$P = 1 - \left(\frac{2}{3}\right)^5 = \frac{211}{243}.$$

(3) 連続 3 回勝つ場合は，1〜3，2〜4，3〜5 回目を勝つ場合の 3 通りあります．よって，

$$P = 3\left(\frac{1}{3}\right)^3\left(\frac{2}{3}\right)^2 = \frac{4}{81}.$$

(4) 後の 2 回は勝っても勝たなくても構いません．よって，

$$P = \left(\frac{1}{3}\right)^3\left(\frac{3}{3}\right)^2 = \frac{1}{27}.$$

(5) 5 回連続して勝つ確率は (1) で求められていて

$$\left(\frac{1}{3}\right)^5 = \frac{1}{243}.$$

4 回連続して勝つ確率は 1〜4 回目勝つまたは 2〜5 回目勝つの 2 通りあり，その確率は

$$2\left(\frac{1}{3}\right)^4\frac{2}{3} = \frac{4}{243}.$$

3 回連続して勝つのは以下の場合です：（ア）1〜3 回目に勝ち，4 回目は勝たず，5 回目は勝っても勝たなくてもよい．（イ）2〜4 回目に勝ち，1 回目と 5 回目は勝たない．（ウ）3〜5 回目に勝ち，2 回目は勝たず，1 回目は勝っても勝たなくてもよい．よって，（ア）〜（ウ）の確率は

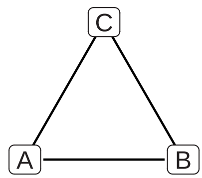
$$\left(\frac{1}{3}\right)^3\frac{2}{3}\times 2 + \left(\frac{1}{3}\right)^3\left(\frac{2}{3}\right)^2 = \frac{16}{243}.$$

これらから求める確率は

$$P = \frac{1+4+16}{243} = \frac{7}{81}$$

となります．なお，2〜4 回目に続けて勝ち，1 回目と 5 回目は勝っても勝たなくてもよい場合を考えると，これは 3 回続けて勝つ場合の一部と，4 回および 5 回続けて勝つ場合を含みます.

15.2.4　確率の漸化式



K 君は白ネズミを飼っています．ネズミを遊ばせようと 3 本の棒で三角形を作りました．その頂点 A，B，C には休憩用の板がついています．ネズミは各頂点から別の 2 つの頂点に等確率で移り，途中で引き返したときは移動しないと考えることにします．K 君はネズミを頂点 A におきました．ネズミが移動を n 回繰り返したとき，頂点 A に戻っている確率 $P_n(\mathrm{A})$ を求めましょう.

ネズミは始め A におかれたので移動前 ($n=0$) には確実に A にいます．このことは $P_0(\mathrm{A}) = 1$ と表されます．1 回目の移動でネズミは B，または C に等確率で移動します．このことを $P_1(\mathrm{B}) = P_1(\mathrm{C}) = \frac{1}{2}$ と表しましょう．一般に，移動を k 回繰り返したときにネズミが頂点 V にいる確率を $P_k(\mathrm{V})$ と書きましょう．よって，例えば 2 回移動したときにネズミが A にいる確率は，A→B→A または A→C→A と移動する場合で

$$P_2(\mathrm{A}) = \frac{1}{2}\cdot\frac{1}{2} + \frac{1}{2}\cdot\frac{1}{2} = P_1(\mathrm{B})\cdot\frac{1}{2} + P_1(\mathrm{C})\cdot\frac{1}{2}$$

と表すことができます．この関係式は後で役立ちます.

ちょっと考えるとわかるように，この問題は今までのやり方で解くことはできません．解くためには数列の漸化式の形にする必要があります．そのためのポイントは 2 つあります．第 1 のポイントは，k 回移動したときにネズミは A，B，C のどこかにいますが，必ずどこかにいますので

$$P_k(\mathrm{A}) + P_k(\mathrm{B}) + P_k(\mathrm{C}) = 1 \qquad (k = 0,\ 1,\ 2, \cdots)$$

が成り立ちます．これを「全確率保存則」といいましょう．第 2 のポイントは，$k+1$ 回移動したときにネズミが A にいるのはその直前に B または C にいる場合で，A に移る確率は共に $\frac{1}{2}$ なので，先の $P_2(\mathrm{A})$ を求めた例にならって

$$P_{k+1}(\mathrm{A}) = (P_k(\mathrm{B}) + P_k(\mathrm{C})) \cdot \frac{1}{2} \qquad (k = 0,\ 1,\ 2, \cdots)$$

が得られます．上式は全確率保存則を用いると，$p_k = P_k(\mathrm{A})$ と簡略して

$$p_{k+1} = \frac{1}{2}(1 - p_k), \quad \text{よって} \quad p_{k+1} = -\frac{1}{2}p_k + \frac{1}{2}$$

と表され，既に学んだ 2 項間漸化式になります．これを初項 $p_0 = 1$ の条件で解きます．

この漸化式を解くのは手頃な練習問題です．各自試みましょう．ヒント：上の漸化式に対する特性方程式

$$\alpha = -\frac{1}{2}\alpha + \frac{1}{2}$$

を考えます．これから $\alpha = \frac{1}{3}$ と定まります．

さて，解きましょう．漸化式から特性方程式を辺々引くと

$$p_{k+1} - \alpha = -\frac{1}{2}(p_k - \alpha)$$

が得られ，数列 $\{p_k - \alpha\}$ が等比数列であることがわかります．これから，

$$p_n - \alpha = \left(-\frac{1}{2}\right)^n (p_0 - \alpha)$$

が得られ，$\alpha = \frac{1}{3}$ と $p_0 = 1$ から最終的に

$$p_n = P_n(\mathrm{A}) = \frac{2}{3}\left(-\frac{1}{2}\right)^n + \frac{1}{3}$$

となります．$P_2(\mathrm{A}) = \frac{1}{2}$ であることを確かめましょう．

では，問題です．いたずら好きの A 坊はおいたが過ぎ，今日はおやつをあげないとお母さんに叱られました．ただし，もうしませんと泣いて頼むので，お母さんはサイコロゲームで A 坊が勝ったときはあげることにしました．ゲームのルールはちょっと複雑です．始め，おやつはお母さんが持っています．A 坊がサイコロを振ります．お母さんからおやつをとるためには 3 以下の目を出

さないといけません．運よくおやつが A 坊に移ってもまだゲームは終わりません．またサイコロを振ります．A 坊がおやつを持っていても，もし 5 または 6 の目が出たときはお母さんに返さないといけません．これを A 坊の年の数 5 だけ繰り返して，つまりサイコロを 5 回振って決着をつけます．そのとき，おやつが A 坊の所にあれば A 坊の勝ちです．A 坊が勝つ確率 $P_5(\mathrm{A})$ を求めなさい．ヒント：おやつは A 坊かお母さんのどちらかの所にあります．練習のために，サイコロを 2 回振ったとき，おやつが A 坊の所にある確率 $P_2(\mathrm{A})$，およびお母さんの所にある確率 $P_2(\text{母})$ を直接求め，$P_2(\mathrm{A}) + P_2(\text{母}) = 1$ を確かめましょう．$P_5(\mathrm{A})$ を場合分けの方法で解くのは面倒です．

サイコロを k 回振ったときに，おやつが K 坊の所にある確率を $P_k(\mathrm{A})$，お母さんの所にある確率を $P_k(\text{母})$ としましょう．まず，確率 $P_2(\mathrm{A})$ については，おやつが 母→A→A または 母→母→A と移る場合で，

$$P_2(\mathrm{A}) = \frac{3}{6}\cdot\frac{4}{6} + \frac{3}{6}\cdot\frac{3}{6} = P_1(\mathrm{A})\cdot\frac{2}{3} + P_1(\text{母})\cdot\frac{1}{2}$$

と表されますね．$P_2(\text{母})$ も同様に計算します．

次に，おやつは A 坊かお母さんのどちらかの所にあるので，全確率保存則

$$P_k(\mathrm{A}) + P_k(\text{母}) = 1 \qquad (k = 0,\ 1,\ 2, \cdots)$$

が成り立ちます．

さて，サイコロを $k+1$ 回振ったときにおやつが A 坊の所にある確率 $P_{k+1}(\mathrm{A})$ は，その直前におやつが A 坊の所にあるかお母さんの所にあるかで分けて

$$P_{k+1}(\mathrm{A}) = P_k(\mathrm{A})\cdot\frac{2}{3} + P_k(\text{母})\cdot\frac{1}{2}$$

が得られます．ここで全確率保存則と簡略記号 $p_k = P_k(\mathrm{A})$ を用いると

$$p_{k+1} = \frac{2}{3}p_k + \frac{1}{2}(1-p_k), \quad \text{よって} \quad p_{k+1} = \frac{1}{6}p_k + \frac{1}{2}$$

が得られます．

この漸化式を条件 $P_0(\text{母}) = 1$，つまり初項 $p_0 = 0$ として解きます．漸化式の特性方程式は $\alpha = \frac{1}{6}\alpha + \frac{1}{2}$ で，$\alpha = \frac{3}{5}$ と決まります．漸化式と特性方程式から

$$p_{k+1} - \alpha = \frac{1}{6}(p_k - \alpha)$$

が得られるので，これを解いて

$$p_n - \alpha = \left(\frac{1}{6}\right)^n (p_0 - \alpha).$$

よって，$\alpha = \dfrac{3}{5}$ と $p_0 = 0$ から

$$p_n = -\frac{3}{5}\left(\frac{1}{6}\right)^n + \frac{3}{5}$$

となります．今の場合，サイコロを 5 回振って決着をつけるので，A 坊がおやつを貰える確率は

$$p_5 = P_5(\mathrm{A}) = -\frac{3}{5}\left(\frac{1}{6}\right)^5 + \frac{3}{5} \fallingdotseq 0.5999 \fallingdotseq 60\ \%$$

です．

15.2.5　連続事象の確率

15.2.5.1　一様分布

雨の日に長さ L の紐（ひも）を張り，雨粒が紐に落ちる点の分布を考えましょう．雨粒は紐にランダム（でたらめ，無作為）に当たるので，雨粒が紐のどこに落ちるのも同程度に確からしいと考えられます．よって，紐の上で長さ l の任意の区間を考えると，雨粒の落ちる数は長さ l のみに比例すると考えられ，これを雨粒の分布は **一様分布** であるといいます．よって，紐に当たる雨粒のうちその区間に落ちる確率 P は

$$P = \frac{l}{L}$$

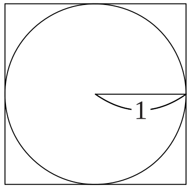
と，長さの比で表されます．

紐が区間 $[0,\ L]$ にあり，長さ l の部分を区間 $[a,\ a+l]$ としましょう．そのとき，雨粒の当たる点を X で表すと，上の確率は

$$P(a \leqq X \leqq a+l) = \int_a^{a+l} p(x)\,dx \qquad \left(p(x) = \frac{1}{L}\right)$$

のように積分を用いて表すことができます．ここで，$p(x)$ は **確率密度** といわれ，今の場合，雨粒が一様に分布することを反映して，定数になります．

今度は一辺が 2 の正方形の板とそれに内接する円の部分を考えます．板に当たる雨粒のうち円の部分に当たる確率を求めましょう．雨粒は板に一様に当たると考えられるので，その確率は円の面積と板の面積の比で表されると考えら



れます．よって，板の位置を $-1 \leqq x \leqq 1$，$-1 \leqq y \leqq 1$，雨粒の当たる点を $(X,\ Y)$ とすると，その確率を

$$P(X^2 + Y^2 \leqq 1) = \frac{\pi 1^2}{2^2} = \frac{\pi}{4}$$

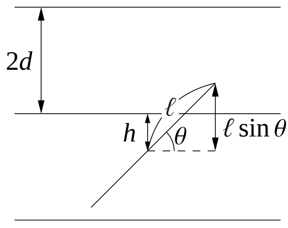
のように表すことができます．

コンピュータでは与えられた区間でランダムに数を発生させるプログラムを作ることができ，そのように作られた数を **乱数** といいます．今，区間 $-1 \leqq x \leqq 1$ と $-1 \leqq y \leqq 1$ で乱数を発生させて，その組 $(x,\ y)$ を一辺が 2 の正方形上の点 $(x,\ y)$ に対応させます．乱数で作られた点の総数を N，そのうち円内にある点の数を n とします．N が十分に大きいとき，n と N の比は円と正方形の面積の比になると考えられるので

$$\frac{n}{N} \fallingdotseq \frac{\pi}{4},\ \text{よって}\quad \pi \fallingdotseq \frac{4n}{N}$$

が成り立ち，これから π の近似値が得られます．例えば，$N = 1000$ のとき π の近似値として 3.12〜3.16 の値が得られます．このように乱数を用いて数学の問題を調べる方法を **モンテカルロ法** といいます．

15.2.5.2　ビュッフォンの針



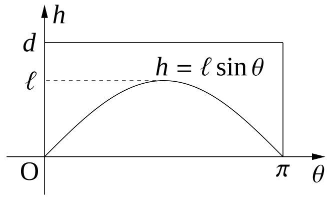
コンピュータがない 18 世紀後半，数学と物理学に熱中したフランスの博物学者ビュッフォンは π の値を求めるのに面白い実験をしました．紙に間隔 $2d$ の平行線を多数引き，その上に長さ 2ℓ $(2\ell < 2d)$ の針を何回も投げます．針が平行線と交わる割合を調べると π の値が求められるというのです．

その原理はこうです．針の中心とその点から最も近い平行線の距離を h $(0 \leqq h \leqq d)$，針と平行線のなす角を θ $(0 \leqq \theta < \pi)$ とすると，針と平行線が交わる条件は，図からわかるように

$$h \leqq \ell \sin\theta$$

です．

針をランダムにたくさん投げたとき，針の中心と向きはランダムになるので，h と θ の値の分布は $0 \leqq h \leqq d$ と $0 \leqq \theta < \pi$ で一様になると考えられます．よって, 右図の θh 平面上の長方形を考えると，観測される点 $(\theta,\ h)$ がその長方形のどこであるかについては，どこも同程度に確からしいと考えられます．したがって，針と平行線が交わる確率は右図の長方形の面積に対する曲線 $h = \ell \sin\theta$ から下の部分の割合で表されます：



$$P(h \leqq \ell \sin\theta) = \frac{\int_0^{\pi} \ell \sin\theta\, d\theta}{\pi d} = \frac{\Big[-\ell\cos\theta\Big]_0^{\pi}}{\pi d} = \frac{2\ell}{\pi d}$$

一方，針を N 回投げてそのうち平行線と n 回交わるとすると，$\dfrac{n}{N}$ の値が近似的に $P(h \leqq \ell \sin\theta)$ に等しくなるので

$$\frac{n}{N} \fallingdotseq \frac{2\ell}{\pi d},\ \text{よって}\quad \pi \fallingdotseq \frac{2\ell N}{nd}$$

から π の近似値が計算できます．19 世紀後半～20 世紀初めまで多くの人がこの実験を試みたようで，$\pi \fallingdotseq 3.13$～3.16 の値が得られています．君たちも試しにやってみましょう．

§15.3　期待値と分散

15.3.1　期待値

15.3.1.1　パスカルの配分方法

この章の始めに紹介したパスカルが受けたギャンブル相談について議論しておきましょう．それは賭けの途中でゲームを止めざるを得ない場合に賭け金をどう配分するかという問題でした．その問題に対するパスカルの解法がフェルマーとの往復書簡に記されています．彼の示した配分方法が，勝負が決着する前に結果を予測して期待される値を求める理論，つまり確率論の始まりとされています．

その配分問題を簡単な例で議論しましょう．ゲームのルールは，A，B の二人が賭け金を 10 万円ずつ出し合って，先に 3 勝したほうが賭け金 20 万円を

総取りするというものです．A が 2 勝，B が 1 勝したところで，当局のがさ入れ（家宅捜索）があって二人はあわてて逃げ出します．途中で止めたときの結果を考慮すると，賭け金の 20 万円を公平に分けるにはどんな割合が適切かというわけです．A が 2 勝，B が 1 勝しているので勝ち数の比 2：1 で分けるのが適当とも考えられます．しかし，どちらかが 1 勝もしていない場合を考えると勝ち星がないほうの取り分がなくなるので，勝ち数の比で分けるのは公平とはいえません．

パスカルは‘賭けを止めずに続行した場合を想定’して，A，B の各々が 3 勝する確率を考え，その確率の比で配分するのが公平であると考えました．このとき，A，B に力量の差はなくどちらが勝つ確率も $\frac{1}{2}$ であるとするのが公平です．A が 2 勝，B が 1 勝しているので，A は残りのゲームを 2 敗せずに勝つ必要があり，B はもう 1 敗もできません．よって，1 回のゲームで A が勝つ事象を A とし，B が勝つ事象を B とすると，A は A または BA で 3 勝し，B は BB で 3 勝します．したがって，A，B の期待される取り分 E_{A}，E_{B} は

$$E_{\mathrm{A}} = 20\,P(\mathsf{A} \cup \mathsf{BA}) = 20\left(\frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2}\right) = 20 \cdot \frac{3}{4} = 15\ \text{万円}$$
$$E_{\mathrm{B}} = 20\,P(\mathsf{BB}) = 20\,\frac{1}{2} \cdot \frac{1}{2} = 20 \cdot \frac{1}{4} = 5\ \text{万円}$$

のように計算され，期待される値は確率を用いて求めることができます．

15.3.1.2　期待値

A 子さんは携帯電話を使いすぎて今月のお小遣いが千円ほど足りなくなりそうです．そこで，子供に甘いお父さんをサイコロ遊びに誘って小遣いをせしめることにしました．1 個のサイコロを振ります．サイコロの目の数を変数 X で表すとき，$X = k$ $(k = 1,\ 2,\ \cdots,\ 6)$ が出たら $200k$ 円頂戴（ちょうだい）というわけです．彼女が期待できる金額はいくらでしょうか．

期待できる金額とは，サイコロ遊びを何度も繰り返したとき，彼女が貰える金額の平均額を意味し，それを $E(200X)$ と表しましょう．目の数 $X = k$ のサイコロ $\boxed{k}$ が出たら $200k$ 円貰えますが，その確率は $P(X = k)$ $(= \frac{1}{6})$ なので，$\boxed{k}$ が出ることを考えて期待できる金額は $200k\,P(X = k)$ 円です．したがって，

$k = 1, 2, \cdots, 6$ の全体に対して期待できる金額は

$$E(200X) = \sum_{k=1}^{6} 200k\,P(X = k) = \sum_{k=1}^{6} 200k\,\frac{1}{6} = 200\,\frac{6 \cdot 7}{2}\,\frac{1}{6} = 700\ \text{円}$$

となります．実際に彼女が振った結果は ⚁ で 400 円貰いました．なお，1 個のサイコロを振ったときに出る目の数 X の平均値 $E(X)$ は，同様に考えて

$$E(X) = \sum_{k=1}^{6} k\,P(X = k) = \sum_{k=1}^{6} k\,\frac{1}{6} = \frac{7}{2} = 3.5$$

となりますね．

一般に，変数 X のとり得る値が $x_1, x_2, \cdots, x_n$ であるとき，それらに対応して確率 $P(X = x_k)$（ただし，$\sum_{k=1}^{n} P(X = x_k) = 1$）が定まるとき，$X$ は確率を伴った変数なので，それを **確率変数** といいます．このとき，$X = x_k$ と $P(X = x_k)$ の関数関係を確率変数 X の **確率分布** といいます．特に，上の例では $P(X = k)$ は定数であり，このような分布を，連続事象の場合と同様，**一様分布** といいます．我々は確率分布を多くの場合に理論的確率の分布として議論しますが，その分布は試行を無限に繰り返して得られる‘統計的確率の分布に一致する’はずであり，今後の議論のためにもこのことを忘れないようにしましょう．

確率変数 X の期待される値，または平均の値については

$$E(X) = \sum_{k=1}^{n} x_k P(X = x_k)$$

を確率変数 X の **期待値**，または **平均値** と定義します（E は expected value（期待値）の頭文字）．

なお，確率変数 X の 1 次式 $aX + b$（a, b は定数）の期待値は

$$E(aX + b) = \sum_{k=1}^{n} (ax_k + b)P(X = x_k)$$

で与えられ，これから直ちに，定理

$$E(aX + b) = aE(X) + b \qquad (a,\ b\ \text{は定数})$$

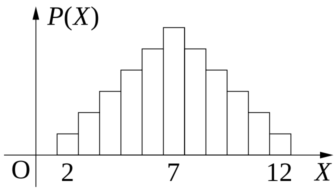
が得られます．各自で証明しましょう．

さて，1000 円にまだ足りない A 子さんは“ ねえ，お父さん．100 円でいいから．”と今度はサイコロを 2 個振りました．出る目の数の和を X として $100X$ 円ほしいというのです．

X を確率変数として確率 $P(X = k)$ の分布を表す表を作ってみましょう．2 つのサイコロを区別すると，根元事象は目の数 1，2，$\cdots$，6 から 2 個とる重複順列の形で表されます：

$$\{(1,1);\ (1,2),(2,1);\ (1,3),(2,2),(3,1);\cdots;(6,6)\}$$

全部で 36 通りありますね．例えば，$X = 2$ または 12 となる場合の数は共に 1 通りしかなく，また，$X = 7$ つまりサイコロの目の数の平均 3.5 の 2 倍となる事象は，最も多くて



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

の 6 通りもあります．他の X の場合については練習問題にしましょう．

確率分布を表にすると

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

となり，$X = 7$ で最大になる山型の分布になります．

A 子さんが期待できる金額は

$$\begin{aligned} E(100X) &= \sum_{k=2}^{12} 100k\,P(X = k) \\ &= \frac{100}{36}(2\cdot 1 + 3\cdot 2 + 4\cdot 3 + 5\cdot 4 + \cdots + 9\cdot 4 + 10\cdot 3 + 11\cdot 2 + 12\cdot 1) \\ &= 700\ \text{円} \end{aligned}$$

となります．A 子さんは ⚂ ⚄ を出してめでたく 800 円をゲットしました．なお，出る目の数の和 X の平均値 $E(X)$ は 1 個のサイコロの場合の平均値 3.5 の 2 倍の 7 で，ここで確率が最も大きくなっていますね．

ところで，2 つのサイコロを振ったときの $E(100X)$ については，それぞれのサイコロの目の数を X_1，X_2 とすると，$X = X_1 + X_2$ と考えることもできます．確率変数 X_1，X_2 の確率分布は $\dfrac{1}{6}$ の一様分布なので，

$$\begin{aligned} E(100X) &= E(100(X_1 + X_2)) = \sum_{k=1}^{6}\sum_{l=1}^{6} 100\,(k+l)P(X_1 = k,\ X_2 = l) \\ &= \sum_{k=1}^{6}\sum_{l=1}^{6} 100\,(k+l)\frac{1}{6^2} = 100\left(\sum_{k=1}^{6} k + \sum_{l=1}^{6} l\right)\frac{1}{6} = 100(3.5+3.5) \\ &= 700\,\text{円} \end{aligned}$$

のようにして計算することもできます．

上の計算方法を一般化すると，確率変数 X，Y と確率分布 $P(X = x_k,\ Y = y_l)$ が与えられるとき，期待値 $E(X+Y)$ についての定理

$$E(X+Y) = E(X) + E(Y) = \sum_k x_k P(X = x_k) + \sum_l y_l P(Y = y_l)$$

が成り立ちます．ここで，$P(X = x_k) = \sum_l P(X = x_k,\ Y = y_l)$ は Y の値を問わない確率です．$P(Y = y_l)$ も同様です．なお，$\sum_k$ は $\sum_{k=1}^{n}$ などを表す簡略記号で，大学ではよく用いられます．上の定理の証明は簡単で

$$\begin{aligned} E(X+Y) &= \sum_k \sum_l (k+l)P(X = x_k,\ Y = y_l) \\ &= \sum_k k \sum_l P(X = x_k,\ Y = y_l) + \sum_l l \sum_k P(X = x_k,\ Y = y_l) \\ &= \sum_k k\,P(X = x_k) + \sum_l l\,P(Y = y_l) = E(X) + E(Y) \end{aligned}$$

で終りです．

なお，サイコロを 2 つ振ったときの確率変数 X は，各サイコロの目の数を X_1，X_2 とすると，$X = X_1 + X_2$ であり，このとき X_1，X_2 の分布は一様分布であるにもかかわらず，X の分布は X_1，X_2 の平均値 3.5 の和 7 をピークとする山型の分布になりましたね．その理由を X_1，X_2 の平均をとる確率変数 $\overline{X}_2 = \dfrac{X_1 + X_2}{2}$ を用いて考えてみましょう．

確率変数 $\overline{X}_2$ は，1 個のサイコロを振る場合と同様，その分布の範囲が $1 \leqq \overline{X}_2 \leqq 6$，平均値 $E(\overline{X}_2)$ が $E(X_1)$，$E(X_2)$ と同様に 3.5 となります．しかし，$\overline{X}_2$ は目の数の平均をとる確率変数なので，得られる $\overline{X}_2$ の値は X_1，X_2 の値と比べて平均値側に寄ったものが多くなります．このことは，君たちのクラスの平均身長を求めるときに，ランダムに 2 人のペアに分けておき，ペアの平均の値をデータとして平均身長を求めることを考えると理解できます．ペアのデータ値は元の各人の値と比べて平均身長により近いものの割合が多いはずです．したがって，平均をとる確率変数 $\overline{X}_2$ は平均値からのばらつきを小さくする働きをもち，$\overline{X}_2$ の確率分布は $\overline{X}_2 = E(\overline{X}_2) = 3.5$ で最大になる山型になると解釈できます．このように考えると，確率変数 $X = X_1 + X_2 = 2\overline{X}_2$ の分布が山型になるのは当然ですね．

もし，n 個 $(n \gg 1)$ のサイコロを振って目の数の平均

$$\overline{X}_n = \frac{X_1 + X_2 + \cdots + X_n}{n} \qquad (n \gg 1)$$

を確率変数にすると，ばらつきを小さくする働きは遙かに強くなり，その確率分布は $\overline{X}_n = E(\overline{X}_n) = E(X_i)$ で最大になる極めて鋭い山型になることが予想され，そのことは後の §で証明されます．さらに，サイコロ振り以外の多くの場合においても，平均をとる確率変数 $\overline{X}_n$ の重要性が確率論における最も重要な 2 つの定理によって強調されるでしょう．それらの定理については後で議論します．

15.3.1.3　期待値の練習問題

練習問題をやりましょう．それぞれ 1，2，3,$\cdots$，n と書いた n 枚のカードがあり，それから無造作に 2 枚を引く．(1) 2 枚のカードに書かれた数字の小さいほうが k である確率を求めよ．(2) 小さいほうの数字の平均値を求めよ．ノーヒントでやってみましょう．どのくらいの値になるか予想してから始めるのがよいでしょう．

$\boxed{1}\ \boxed{2}\ \boxed{3} \cdots \boxed{n}$

(1) は全事象の場合の数が，n 枚から 2 枚を引く組合せなので，${}_n\mathrm{C}_2$ 通りあります．また，小さいほうの数字 X が k である場合の数は，大きいほうの数が $k+1$，$k+2, \cdots$，n のどれかなので，$n-k$ 通りあります．よって，$X = k$ とな

る確率は

$$P(X=k)=\frac{n-k}{{}_n\mathrm{C}_2}=\frac{2(n-k)}{n(n-1)}$$

です．(2) は (1) より

$$E(X)=\sum_{k=1}^{n-1}k\frac{2(n-k)}{n(n-1)}=\frac{2}{n(n-1)}\left\{n\frac{n(n-1)}{2}-\frac{(n-1)n(2n-1)}{6}\right\}=\frac{n+1}{3}$$

となります．予想は当たりましたか．

期待値の問題には Σ 計算の技術を要するものが多くあります．1 題やってみましょう．A，B の二人がジャンケンをして A が勝ったら終了する．ジャンケンの平均回数は何回だろうか．ただし，二人のどちらかが上手ということはない．君は何回と予想するかな？ Σ 計算で微分の公式

$$\sum_{k=1}^{n}kx^{k-1}=\frac{d}{dx}\sum_{k=1}^{n}x^k$$

は役に立ちます．

1 回のジャンケンで A が勝つ，負ける，引き分ける確率は全て $\frac{1}{3}$ です．よって，ジャンケンの回数を X として，k 回目に初めて A が勝つ確率は

$$P(X=k)=\left(\frac{2}{3}\right)^{k-1}\left(\frac{1}{3}\right)$$

よって，ジャンケンの平均回数 $E(X)$ は，いつまで経っても A が勝てない場合もあるので，$p=\frac{2}{3}$ とおくと

$$E(X)=\lim_{n\to\infty}\sum_{k=1}^{n}kp^{k-1}(1-p)$$

となります．これを計算するのに先の公式を用いると

$$\begin{aligned}E(X)&=\lim_{n\to\infty}(1-p)\frac{d}{dp}\sum_{k=1}^{n}p^k=\lim_{n\to\infty}(1-p)\frac{d}{dp}\frac{p-p^{n+1}}{1-p}\\&=\lim_{n\to\infty}\left(1-(n+1)p^n+\frac{p-p^{n+1}}{1-p}\right)\end{aligned}$$

となります．ここで，$p=\frac{2}{3}$ より，$n\to\infty$ のとき $p^n\to 0$，$np^n\to 0$ なので

$$E(X)=1+\frac{p}{1-p}=\frac{1}{1-p}=3\text{ 回}$$

が得られます．3 回やれば平均して 1 回は勝つのだから当然の結果ですね．

15.3.2　分散と標準偏差

15.3.2.1　期待値からのずれ

前の §§で，A 子さんはお父さんにお小遣いをねだってサイコロを振りましたね．1 個のサイコロを振ったとき期待できる金額は 700 円でした．しかし，1 個のサイコロの目の数 X はどれも同程度に出やすいので，平均的には 700 円だとしても，実際に振ったときに貰える金額はばらつきが大きく，平均額の 700 円からずれる場合が多いと思われます．このような場合には，期待値は実際に貰える金額のよい目安であるとはいえないでしょう．一方，サイコロを 2 個振った場合には，目の数の和 X に対する確率分布は 7 で最大で，7 から離れると小さくなります．よって，X の平均が 7 になったことは，実際にサイコロを振ったときに 7 から大きくずれた X の値が出ることは多くないことを意味します．

このように，平均値を実際の試行で得られる値の目安として考えるときには，確率分布のばらつきの度合，よって平均値からずれる度合は大きな問題になります．

そこで，一般の確率変数 X に対して，実際の試行で得られる X の値のばらつきを数値的に表現することを考えましょう．それにはそれぞれの X の値とその期待値 $E(X)$ との差の 2 乗の平均値を表す X の **分散** といわれる量 $\overline{(X-E(X))^2}$ が便利であり，それを記号 $V(X)$ で表しましょう．具体的には，X が個々にとり得る値を x_k $(k=1,\ 2,\cdots,n)$，X の確率分布を $P(X=x_k)$ と書くと，X の分散は

$$V(X) = \overline{(X-E(X))^2} = \sum_k (x_k - m)^2 P(X=x_k) \qquad （ただし，m = E(X)）$$

と定義されます．$V(X)$ の V は variance（分散）の頭文字，また m は mean value（平均値）の頭文字です．

X のばらつきを考えるには $(X-E(X))^2$ より $|X-E(X)|$ のほうが適しているように思えますので，$|X-E(X)|$ に当たる量として **標準偏差** といわれる

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

が用意されています．$\overset{シグマ}{\sigma}$ は standard deviation（標準偏差）の頭文字 s に当た

るギリシャ文字で，Σ の小文字です．標準偏差はばらつきに関する議論で直接重要になる量であり，実際の試行で得られる X の値は，その大半が

$$|X - E(X)| < \sigma(X) \iff E(X) - \sigma(X) < X < E(X) + \sigma(X)$$

を満たすものといってよいでしょう．もし $\sigma(X) = 0$ ならば，ばらつきはまったくなく，$P(X = E(X)) = 1$ です．標準偏差 $\sigma(X)$ の詳しい議論は後で行うことにしましょう．なお，標準偏差は君たちが大嫌いな「偏差値」に直接関係しています．

練習として，1 個のサイコロを振ったときに出る目の数 X の分散 $V(X)$ と標準偏差 $\sigma(X)$ を計算してみましょう．X の平均値は $E(X) = m\ (= \frac{7}{2})$ ですから

$$\begin{aligned} V(X) &= \sum_{k=1}^{6}(k-m)^2 P(X=k) = \sum_{k=1}^{6}(k^2 - 2mk + m^2)P(X=k) \\ &= \sum_{k=1}^{6} k^2 P(X=k) - 2m\sum_{k=1}^{6} kP(X=k) + m^2\sum_{k=1}^{6} P(X=k) \\ &= E(X^2) - 2mE(X) + m^2 = E(X^2) - \{E(X)\}^2 \end{aligned}$$

と変形しておくと

$$E(X^2) = \sum_{k=1}^{6} k^2 P(X=k) = \frac{6 \cdot 7 \cdot 13}{6}\frac{1}{6} = \frac{7 \cdot 13}{6}$$

より

$$\begin{aligned} V(X) &= E(X^2) - \{E(X)\}^2 = \frac{7 \cdot 13}{6} - \left(\frac{7}{2}\right)^2 = \frac{35}{12}, \\ \sigma(X) &= \sqrt{V(X)} = \sqrt{\frac{35}{12}} \fallingdotseq 1.7 \end{aligned}$$

が得られます．よって，出る目の数 X の大半は $E(X)-\sigma(X) < X < E(X)+\sigma(X)$ より $3.5 - 1.7 = 1.8 < X < 5.2 = 3.5 + 1.7$，つまり 2〜5 の目です．当たり前ですね．これは X の確率分布が一様なために標準偏差が大きくなったからです．$2 \leqq X \leqq 5$ となる確率は $\frac{4}{6} \fallingdotseq 67\ \%$ を占めます．このことを，$m = E(X)$, $\sigma = \sigma(X)$ と略記して，$P(|X - m| < \sigma) = \frac{4}{6}$ のように表しましょう．

上の計算で関係式

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

を用いましたが，導出の過程で確率の一般的性質だけを用いていることからわかるように，これは分散に関する一般公式です．同様に，確率変数 X の 1 次式 $aX + b$ の分散について，公式

$$V(aX + b) = a^2 V(X) \qquad （a,\ b \text{ は定数}）$$

が成り立ちます．証明は簡単なので，各自でやってみましょう．

次に，2 個のサイコロを振ったときに出た目の数の和 X の分散を計算しましょう．X の確率分布を直接用いるとただの計算になってしまうので，各々のサイコロの目の数を X_1，X_2 として $X = X_1 + X_2$ を用いましょう．まず，

$$V(X) = V(X_1 + X_2) = E((X_1 + X_2)^2) - \{E(X_1 + X_2)\}^2$$

において，$E(X_1 + X_2) = E(X_1) + E(X_2)$ だから，$E(X_1) = m_1$，$E(X_2) = m_2$ とおくと

$$\begin{aligned} V(X) &= E(X_1^2 + 2X_1X_2 + X_2^2) - (m_1^2 + 2m_1m_2 + m_2^2) \\ &= E(X_1^2) - m_1^2 + E(X_2^2) - m_2^2 + 2(E(X_1X_2) - m_1m_2) \\ &= V(X_1) + V(X_2) + 2(E(X_1X_2) - m_1m_2) \end{aligned}$$

となります．ここまでは期待値と分散の一般公式を用いて導かれます．

ここで，サイコロ振りの試行が独立試行であることを用いて，定理

$$E(X_1X_2) = E(X_1) \cdot E(X_2) = m_1m_2$$

が成り立つことを導きましょう．各々のサイコロの目の数 X_1，X_2 がそれぞれ k，l である事象を $X_1 = k$，$X_2 = l$ とすると，それらの事象は独立であるので $X_1 = k \cap X_2 = l$ が起こる確率について

$$P(X_1 = k \cap X_2 = l) = P(X_1 = k) \cdot P(X_2 = l)$$

が成り立ちますね．このとき，確率変数 X_1，X_2 は **独立** であるといいます．ただし，$P(X_1 = k)$ は X_2 の値を問わない確率，$P(X_2 = l)$ は X_1 の値を問わない確

率なので

$$P(X_1 = k) = \sum_{l=1}^{6} P(X_1 = k \cap X_2 = l), \qquad P(X_2 = l) = \sum_{k=1}^{6} P(X_1 = k \cap X_2 = l)$$

です．よって，

$$\begin{aligned} E(X_1X_2) &= \sum_{k=1}^{6}\sum_{l=1}^{6} kl P(X_1 = k \cap X_2 = l) = \sum_{k=1}^{6}\sum_{l=1}^{6} kl P(X_1 = k)P(X_2 = l) \\ &= \sum_{k=1}^{6} kP(X_1 = k)\left(\sum_{l=1}^{6} lP(X_2 = l)\right) = E(X_1)E(X_2) = m_1m_2 \end{aligned}$$

つまり，$E(X_1X_2) = E(X_1)E(X_2) = m_1m_2$ が成り立ちます．

以上のことから

$$V(X) = V(X_1 + X_2) = V(X_1) + V(X_2)$$

が成り立ち，1 個のサイコロを振る場合に得られた結果 $V(X_1) = V(X_2) = \dfrac{35}{12}$ より

$$V(X) = \frac{35}{6}, \qquad \sigma(X) = \sqrt{\frac{35}{6}} \fallingdotseq 2.4$$

が得られます．

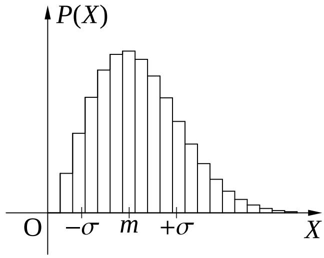
X の平均値は $E(X) = m = 7$ だから，目の数の和 $X = 2$〜12 のうちで，出るものの大半は，$7-2.4 = 4.6 < X < 9.4 = 7+2.4$ より，5〜9 ですね．$5 \leqq X \leqq 9$ である確率は $\dfrac{4+5+6+5+4}{36} = \dfrac{2}{3} \fallingdotseq 67$ % あります．このことは 1 個のサイコロを振ったときの例にならって，$P(|X-m| < \sigma) = \dfrac{2}{3}$ と表されますね．

上で行った式変形のやり方からわかるように，一般に確率変数 X，Y が**独立**であるとき

$$E(XY) = E(X) \cdot E(Y), \qquad V(X+Y) = V(X) + V(Y)$$

が成り立ちます．その証明は上で行ったものとほとんど同じなので，各自試みましょう．その際，X，Y のとる値を x_k，y_l 等として，和については簡略記号 $\sum\limits_k$，$\sum\limits_l$ を用いるのが便利です．

15.3.2.2　標準偏差に関する不等式



確率変数 X の値のばらつきを調べるのに標準偏差 $\sigma(X)$ が有効でした．実際の試行で得られる X の値の大半は，期待値を $m = E(X)$，標準偏差を $\sigma = \sigma(X)$ として，$|X - m| < \sigma$ の範囲にあり，それから外れることは少ないと述べました．右図は確率分布の代表例を表し，$X = m \pm \sigma$ の位置は $\pm\sigma$ で示されています．それに関して「チェビシェフの不等式」といわれる確率分布と標準偏差の関係を表す定理があります：

$$P(|X - m| \geqq |z\sigma|) \leqq \frac{1}{z^2} \qquad (z \text{ は } 0 \text{ でない実数}).$$

この不等式は確率変数 X の分布に一般的な制限を与え，X の値が平均値 m から $|z\sigma|$ 以上離れている部分全体の確率は $\frac{1}{z^2}$ 以下になることを意味します．全確率は 1 ですから，この不等式は不等式

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

と同じです．

この不等式を示しましょう．確率変数 X のとる値を x_1，x_2，$\cdots$，x_n，対応する確率分布を p_1，p_2，$\cdots$，p_n と略記すると，分散 $V(X)$ は

$$\sigma^2 = E((X - m)^2) = \sum_k (x_k - m)^2 p_k$$

ですが，ここで k についての和で $|X - m| < |z\sigma|$ である部分の和をとる記号を

$$\sum_{|x_k - m| < |z\sigma|}$$

等で表すと

$$\begin{aligned}
\sigma^2 &= \sum_k (x_k - m)^2 p_k = \sum_{|x_k - m| < |z\sigma|} (x_k - m)^2 p_k + \sum_{|x_k - m| \geqq |z\sigma|} (x_k - m)^2 p_k \\
&\geqq \sum_{|x_k - m| \geqq |z\sigma|} (x_k - m)^2 p_k \geqq \sum_{|x_k - m| \geqq |z\sigma|} (z\sigma)^2 p_k = z^2\sigma^2 \sum_{|x_k - m| \geqq |z\sigma|} p_k \\
&= z^2\sigma^2 P(|X - m| \geqq |z\sigma|)
\end{aligned}$$

が得られ，したがって，

$$\sigma^2 \geqq z^2\sigma^2 P(|X-m| \geqq |z\sigma|) \quad \Leftrightarrow \quad P(|X-m| \geqq |z\sigma|) \leqq \frac{1}{z^2}$$

が示されます．この不等式は X のとり得る値 x_1，x_2，$\cdots$，x_n の個数が多いときに重要になります．

§15.4　二項分布

15.4.1　サイコロ振りと統計的確率

1 個のサイコロを振る試行を何度も繰り返して ⚀ が出る回数 X の平均値 $E(X)$ とその標準偏差 $\sigma(X)$ を調べましょう（サイコロを何度も振る試行は多くのサイコロを同時に振る試行と考えても構いません）．振る回数を n とすると，6 回に 1 回は ⚀ が出るはずだから，$E(X) = \dfrac{n}{6}$ が予想されます．また，実際の試行で得られる X の値のばらつきは平均値 $E(X)$ からそれほどずれないはずなので，回数 n が大きいときは $\sigma(X)$ は n に比べて非常に小さいことが予想されます．特に，サイコロが正しく作られていれば，$n \to \infty$ のとき，⚀ が出る比率が平均して 6 回に 1 回の割合からずれることはなくなると予想されます．そのことは，サイコロ振りの理論的確率と統計的確率が一致することの証でもあります．このことを検証しましょう．

まず，X の確率分布を求めて真面目（まじめ）に平均値 $E(X)$ を計算してみましょう．サイコロ振りの反復試行は独立試行です．よって，§§15.2.3.2 で議論したように，n 回振ったときに ⚀ が k 回出る確率 $P(X=k)$ は，1，2，3,$\cdots$，n 回から ⚀ が出る k 回を選ぶ方法の総数が ${}_n\mathrm{C}_k$ なので

$$P(X=k) = {}_n\mathrm{C}_k\left(\frac{1}{6}\right)^k\left(\frac{5}{6}\right)^{n-k}$$

となります．

よって，⚀ が出る平均回数 $E(X)$ は，簡単のため $p = \dfrac{1}{6}$ とおくと

$$E(X) = \sum_{k=0}^{n} kP(X=k) = \sum_{k=0}^{n} k\,{}_n\mathrm{C}_k p^k(1-p)^{n-k}$$

で与えられます．これを計算するには §§15.1.2.5 で学んだ公式

$$np = \sum_{k=0}^{n} k\,{}_n\mathrm{C}_k p^k (1-p)^{n-k}$$

が必要で，これを用いると直ちに

$$E(X) = np = \frac{n}{6}$$

が得られます．予想通り，平均して 6 回に 1 回は ⚀ が出ますね．

以上が真面目な計算です．${}_n\mathrm{C}_k$ を巻き込む Σ 計算が必要でしたね．この計算法だと標準偏差の計算はもっと面倒になります．そこで，標準偏差の計算にも使える要領のよい方法でやり直しましょう．

サイコロを振る i 回目の試行に対して，確率変数 X_i を考え，⚀ が出たら $X_i = 1$，⚀ が出なかったら $X_i = 0$ とします．すると，n 回の試行に対して ⚀ が出る回数 X は

$$X = X_1 + X_2 + X_3 + \cdots + X_n$$

となります．このとき，

$$E(X) = E(X_1 + X_2 + X_3 + \cdots + X_n) = \sum_{i=1}^{n} E(X_i)$$

が成り立ちますが，$E(X_i)$ の定義によって

$$E(X_i) = 1 \cdot p + 0 \cdot (1-p) = p$$

ですから

$$E(X) = \sum_{i=1}^{n} p = np = \frac{n}{6}$$

が得られます．これはうまい手ですね．

次に，X の分散 $V(X)$ を求めて標準偏差 $\sigma(X)$ を議論しましょう．上の方法は分散についても適用できます．サイコロを振る反復試行は独立試行なので，確率変数 X_i，X_j $(i \neq j)$ は独立であり

$$V(X) = V(X_1 + X_2 + X_3 + \cdots + X_n) = \sum_{i=1}^{n} V(X_i)$$

が成り立ちます．よって，

$$V(X_i) = E(X_i^2) - \{E(X_i)\}^2 = \{1^2 p + 0^2(1-p)\} - \{p\}^2 = p(1-p)$$

より

$$V(X) = \sum_{i=1}^{n} p(1-p) = np(1-p), \qquad \sigma(X) = \sqrt{np(1-p)}$$

が得られます．

標準偏差 $\sigma(X)$ が $\sqrt{n}$ に比例して増加するので，1 回当たりの試行に換算して起こる回数（相対度数）を表す確率変数，つまり X_1，$X_2, \cdots$，X_n の平均をとる確率変数

$$\overline{X} = \frac{X}{n} = \frac{X_1 + X_2 + X_3 + \cdots + X_n}{n}$$

を考えましょう（平均をとる確率変数 $\overline{X}$ の一般的議論のために，この試行は n 個のサイコロを同時に振る試行と同じだと意識しておきましょう）．すると，

$$E(\overline{X}) = E\left(\frac{X}{n}\right) = \frac{1}{n}E(X) = p = \frac{1}{6}$$

となり，$E(\overline{X}) = \dfrac{1}{6}$ は平均して 6 回に 1 回の割合で ⚀ が出ることを直接表します．

このとき，標準偏差 $\sigma(\overline{X})$ については，$V(aX) = a^2V(X)$，したがって，$\sigma(aX) = |a|\sigma(X)$ より

$$\begin{aligned}\sigma(\overline{X}) &= \sigma\left(\frac{X}{n}\right) = \frac{1}{n}\sigma(X) = \frac{1}{n}\sqrt{np(1-p)} \\ &= \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{5}{36n}}\end{aligned}$$

となります．よって，$\sigma(\overline{X})$ は $\sqrt{n}$ に反比例するので，$n \to \infty$ のとき $\sigma(\overline{X}) \to 0$ となります．このことは，サイコロを多く振れば振るほど，⚀ が出る比率は $\dfrac{1}{6}$ に近づいていき，それからずれることはなくなる，つまり統計的確率は $\dfrac{1}{6}$ であることを意味します．以上の議論は，もちろん，サイコロの他の目についても同様に成り立ちます．

このような収束は次の §§ で議論する「大数(たいすう)の法則」として知られており，統計的確率と理論的確率の一致を保証しています．

15.4.2　二項分布と大数の法則

前の §§のサイコロ振りの反復試行の議論はそのまま一般化できます．ある試行において，事象 A の起こる確率を p とし，その試行を n 回繰り返します．そのとき，事象 A の起こる回数を確率変数 X とすれば，A が k 回起こる確率は

$$P(X = k) = {}_n\mathrm{C}_k p^k (1-p)^{n-k} \qquad (k = 0,\ 1,\ 2, \cdots,\ n)$$

で表されます．このような確率の式で表される確率分布は特に有用な分布なので，これを **二項分布** (binomial distribution) といい，記号 $B(n,\ p)$ で表します．

二項分布は視聴率や世論等の統計調査，製品の抜きとり検査など，非常に大きな集団からサンプルをランダムに抽出する場合などにも現れます．集団の中である性質 A をもつものの比率が p であるとき，集団から n 個を選んでとり出します．性質 A をもつものの個数を X として $X = k$ である確率は，1 個ずつ順に n 個抽出して戻さないと考えると，p が 0 または 1 に極めて近い場合を除いて，よい近似で二項分布 $B(n,\ p)$ の確率分布 ${}_n\mathrm{C}_k p^k (1-p)^{n-k}$ になりますね．

$B(n,\ p)$ に対して，確率変数 X $(= 1,\ 2,\ \cdots,\ n)$ の期待値と標準偏差は，サイコロ振りの反復試行の場合と同様にして

$$E(X) = np, \qquad \sigma(X) = \sqrt{np(1-p)}$$

で与えられます．また，相対度数を表す，または平均をとる確率変数 $\overline{X} = \dfrac{X}{n}$ の期待値と標準偏差は

$$E(\overline{X}) = p, \qquad \sigma(\overline{X}) = \sqrt{\frac{p(1-p)}{n}}$$

となります．

さて，理論的確率と統計的確率は一致すると述べました．そのことを上の結果と一般の場合に成り立つチェビシェフの不等式 $P(|X - m| \geqq |z\sigma|) \leqq \dfrac{1}{z^2}$ を用いて議論しましょう．チェビシェフの不等式の X は一般の確率変数でよいので，X を $\overline{X}$ に置き換え，$\sigma(\overline{X})$ を $\overline{\sigma}$ と書くと

$$P(|\overline{X} - p| \geqq |z\overline{\sigma}|) \leqq \frac{1}{z^2}$$

となります．ここで，$|z\overline{\sigma}| = \varepsilon\ (> 0)$ とおくと

$$P(|\overline{X} - p| \geqq \varepsilon) \leqq \left(\frac{\overline{\sigma}}{\varepsilon}\right)^2$$

が得られます．全確率は 1 なので，この不等式は

$$P(\left|\overline{X}-p\right|<\varepsilon) \geqq 1-\left(\frac{\overline{\sigma}}{\varepsilon}\right)^2$$

と同じです．ここで，$n\to\infty$ のとき $\overline{\sigma}=\sigma(\overline{X})\to 0$ なので，任意の $\varepsilon\ (>0)$ に対して

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

が成り立ちます．

　これは，相対度数を表す，または平均をとる確率変数 $\overline{X}=\dfrac{X}{n}$ の平均値が p であるとき，試行回数 n またはサンプル数 n を大きくしていくと，平均をとる確率変数 $\overline{X}$ の得られる値のばらつき $\sigma(\overline{X})$ は小さくなっていき，$n\to\infty$ の極限では平均値 p からずれることはまったくなくなることを意味します．この定理は（ベルヌーイの）「大数の法則」といわれ，サイコロを何回も振るとき特定の目が 6 回に 1 回の割合で出ることを保証します．この法則の応用範囲は広く，例えば保険会社は，加入者数に対する死亡者数や事故者数が精確に予測可能であるので，保険料を算定する根拠として用います．また，この法則からわかることは，宝くじをたくさん買っても儲かることは少なく，うまく当てたら後は買わずに勝ち逃げするのが利口だということです．

　なお，大数の法則は確率分布が二項分布でなくとも導くことができます．n 個の確率変数 X_1，X_2，$\cdots$，X_n を考えます．それらは独立な確率変数であり，また X_i の期待値 $E(X_i)$，標準偏差 $\sigma(X_i)$ はそれぞれ i によらない共通の値 m, σ であるとします．それらの確率変数は互いに影響しない同種のものなら何でもよいのですが，例えばサイコロを n 回振るとき i 回目に出た目の数を X_i とするとか，視聴率の調査で n 人をランダムに抽出するとき i 人目の人が人気の歌番組を見ていたかいないかで X_i を 1, 0 とするなどと思えばよいでしょう．

　このとき，平均をとる確率変数

$$\overline{X}=\frac{X_1+X_2+\cdots+X_n}{n}$$

の期待値と標準偏差は

$$E(\overline{X})=m, \qquad \sigma(\overline{X})=\frac{\sigma}{\sqrt{n}} \qquad (m=E(X_i), \quad \sigma=\sigma(X_i))$$

となります．これを確かめるのは練習問題としましょう．

上式の 2 つの標準偏差 $\sigma(\overline{X})$ と σ の関係は重要です．例えば，n 個のサイコロを振る例でいうと，$\overline{X}$ は出た目の数の平均をとる確率変数なので，$\overline{X}$ のばらつきは 1 個のサイコロを振る場合の X_i のばらつきに比べて $\dfrac{1}{\sqrt{n}}$ に減ることを意味します．このようなばらつきの減少は §§15.3.1.2 のサイコロを 2 個振る場合の議論から予想されていましたね．このとき，大雑把にいうと，確率分布 $P(\overline{X})$ は個々の分布 $P(X_i)$ を $\overline{X} = m$ の周りに $\dfrac{1}{\sqrt{n}}$ 倍だけ圧縮したようなものになります．

一方，チェビシェフの不等式は

$$P(\left|\overline{X} - m\right| \geqq \left|z\sigma(\overline{X})\right|) \leqq \frac{1}{z^2}$$

となるので，$\varepsilon = \left|z\sigma(\overline{X})\right| = \left|z\dfrac{\sigma}{\sqrt{n}}\right|$ とおくと，

$$P(\left|\overline{X} - m\right| \geqq \varepsilon) \leqq \left(\frac{\sigma}{\varepsilon\sqrt{n}}\right)^2 \Leftrightarrow\ P(\left|\overline{X} - m\right| < \varepsilon) \geqq 1 - \left(\frac{\sigma}{\varepsilon\sqrt{n}}\right)^2$$

となります．これから，$n \to \infty$ のとき，任意の $\varepsilon\ (> 0)$ に対して

$$\lim_{n\to\infty} P(\left|\overline{X} - m\right| < \varepsilon) = 1 \qquad \text{（大数の（弱）法則）}$$

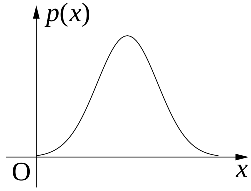
が成り立つことが導かれます（確かめましょう）．これは「大数の（弱）法則」といわれる定理で，平均をとる確率変数 $\overline{X} = \dfrac{X_1 + X_2 + \cdots + X_n}{n}$ の確率分布 $P(\overline{X})$ は $n \to \infty$ のとき $\overline{X}$ の平均値 m に集中することを意味します．サイコロの例でいうと，n 個のサイコロを同時に振るとき，各サイコロの目の数の平均 $\overline{X}$ の確率分布 $P(\overline{X})$ は，$n \to \infty$ のとき $\overline{X}$ の平均値 $m = 3.5$ に集中し，それからずれることはなくなります．また，視聴率の例でいうと，サンプル数 n が非常に多くなると，ある調査で得られた視聴率が，非常に小さな誤差の範囲内で，真の視聴率を反映することを意味します．

§15.5　正規分布

サイコロを振る回数や視聴率のサンプル数が多くなると，確率分布関数はそのままでは扱いが難しくなります．扱い易くなるように，分布関数の近似を考えましょう．

15.5.1　離散分布から連続分布へ

的に向かって矢を射るとき，的の中心から矢が当たった位置までの距離を確率変数 X とすると，事象 $X = x$ $(\geqq 0)$ は連続事象です．このとき，1 点 $X = x$ となる確率は事実上 0 ですから，範囲 $a \leqq X < b$ などで矢が当たる確率 $P(a \leqq X < b)$ を考えることになります．$P(a \leqq X < b)$ は，§§15.2.5 で議論したように，確率密度 $p(x)$ を用いて



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

と表すのが便利です．連続的に分布する確率密度 $p(x)$ を知るためには，矢を無数に射なければなりませんね．このとき，$p(x)$ はどんな曲線になるのでしょうか．

次の例として，ある人気テレビ番組の視聴率を考えましょう．視聴率は，テレビを所有する全世帯のうち，何世帯がその番組を見ていたかを示す比率 p です．全世帯を調べることはできないので，そのうちの n 世帯をサンプル[2)]に選んで調べます．その番組を見ていた世帯数を確率変数 X とすると，ある調査で $X = k_s$ 世帯が見ていたなら視聴率 p は近似的に $\dfrac{k_s}{n}$ であると考えられます．ただし，サンプル数 n が小さいときは信用できないので，n をどのくらい大きくとれば十分であるか，また，結果のばらつき，つまり誤差がどのくらいあるかを考える必要があります．ばらつきを調べるためには，サンプルを何度もとり直して X の標準偏差を調べるようなことが必要になります．よって，確率を考慮した扱いが必要になります．

テレビを所有する世帯数 N は非常に多いので，n 世帯 $(n \ll N)$ のサンプルのうち $X = k$ 世帯がその番組を見ていた確率は，先に議論したように，二項分布：$P(X = k) = {}_n\mathrm{C}_k p^k (1-p)^{n-k}$ $(k = 0,\ 1,\ 2, \cdots, n)$ で表されますね．近似的視聴率 $\dfrac{k_s}{n}$ やその誤差を正しく評価するためにはサンプル数 n をできる限り大きくする必要があります．それは二項分布の式が非常に大きい n に対してどの

2) 用語 **サンプル**（標本）は統計学の用語です．サンプルを抜きとる集団を **母集団** といい，今の場合テレビを所有する全世帯を指します．§§15.1.1 のコイン投げの試行で根元事象の集合（全事象）をサンプル空間（標本空間）といいましたが，各サンプル世帯の視聴状態を根元事象と考えるとサンプル空間とサンプル全体は同質のものといえます．

ように近似されるかを調べればわかります．それについては次の §§で議論しましょう．

15.5.2　正規分布の導出

n が非常に大きいとき，二項分布 $B(n,\ p)$：

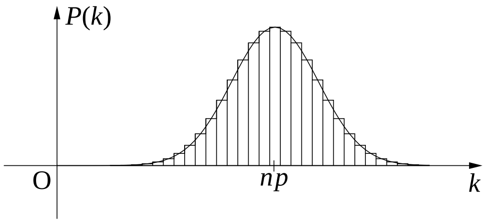
$$P(X=k)={}_n\mathrm{C}_k p^k q^{n-k} \quad (k=0,\ 1,\ 2,\cdots,\ n),$$
$$E(X)=np, \quad V(X)=npq \quad (\text{ただし，} q=1-p)$$

がどのように近似されるかを調べましょう．正しくいうと，np と nq が共に大きいときの近似です．大数の法則によって，確率分布 $P(X=k)$（以下 $P(k)$ と略記）は X の平均値 np 付近に集中します．よって，平均値付近においてはよい近似となり，そこから遠いところでは消え失せるような $P(k)$ を求めることになります．

近似式を導くために，まず二項係数 ${}_n\mathrm{C}_k=\dfrac{n!}{k!(n-k)!}$ に対して，§§14.8.3 で学んだ解析的階乗関数

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

を利用しましょう（n は自然数，x は実数）．この関数を利用すると，${}_n\mathrm{C}_k$ の k は実数に拡張でき，${}_n\mathrm{C}_k$ は k の実数関数と見なせます．よって，二項分布 $P(k)$ も k の実数関数に拡張できます．図はその模式的な様子を表しています．



$P(k)={}_n\mathrm{C}_k p^k q^{n-k}$ は，$E(X)=np$ および大数の法則によって $k=np$ 付近に集中し，そこで $P(k)$ は最大になることが予想されます．よって，$k=np$ 付近における $P(k)$ の近似式を求めましょう．そのために，$P(k)$ の対数をとり，実数関数

$$f(k)=\log P(k)=\log {}_n\mathrm{C}_k+k\log p+(n-k)\log q$$

を考えます．そのために §§13.2.2 で学んだテイラーの定理：

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k + \frac{f^{(n+1)}(a+\theta(x-a))}{(n+1)!}(x-a)^{n+1} \qquad (0 < \theta < 1)$$

を用いて，$k = np$ 付近における $f(k)$ の近似式を求めましょう．

$P(k)$ は $k = np$ で最大かつ極大，よって $f'(np) = 0$ となることが予想されるので，近似式は k の 2 次まで求める必要があります：

$$f(k) \fallingdotseq f(np) + f'(np)(k-np) + \frac{f''(np)}{2!}(k-np)^2.$$

$f'(k)$ を求めるためには，微分の公式を用いるか，または

$$\{\log {}_n\mathrm{C}_k\}' = \lim_{\varDelta k \to 0} \frac{\log {}_n\mathrm{C}_{k+\varDelta k} - \log {}_n\mathrm{C}_k}{\varDelta k}$$

のように平均変化率を考えて極限操作 $\varDelta k \to 0$ をしなければいけません．求めるのは近似式ですから $\varDelta k \to 0$ なしで済ませることを考えましょう．

そのために，1 次式 $ak + b$ については，任意の増分 $\varDelta k$ に対して，微分係数は平均変化率に一致する，つまり

$$\{ak + b\}' = \frac{a(k + \varDelta k) + b - (ak + b)}{\varDelta k} = a$$

が成り立つことに注意しましょう．関数 $y = \log x$ のグラフは，x が十分に大きいところでは，直線のように見えます．${}_n\mathrm{C}_k$ は，二項係数の性質より，n および $k \fallingdotseq np$，$n - k \fallingdotseq nq$ が共に大きいところで大きな値をとり，よって $\log {}_n\mathrm{C}_k$ は $k = np$ 付近の小さな区間においては k の 1 次式に近い振る舞いをします．よって，その導関数は $\varDelta k = 1$ とする平均変化率で近似でき

$$\begin{aligned}\{\log {}_n\mathrm{C}_k\}' &\fallingdotseq \frac{\log {}_n\mathrm{C}_{k+\varDelta k} - \log {}_n\mathrm{C}_k}{\varDelta k} = \log \frac{{}_n\mathrm{C}_{k+1}}{{}_n\mathrm{C}_k} = \log \frac{k!(n-k)!}{(k+1)!(n-k-1)!} \\ &= \log \frac{n-k}{k+1} \fallingdotseq \log \frac{n-k}{k}\end{aligned}$$

となります．したがって，

$$f'(np) \fallingdotseq \log \frac{n - np}{np} + \log p - \log q = \log \frac{nqp}{npq} = \log 1 = 0$$

となります．$f'(np) \fallingdotseq 0$ の結果は $P(k)$ が $k \fallingdotseq np$ で最大になることを反映し，また微分係数を平均変化率で近似したことが適切であることを示しています．

$k = np$ 付近で

$$f''(k) = \{\log {}_n\mathrm{C}_k\}'' \fallingdotseq \frac{-1}{n-k} - \frac{1}{k} = -\frac{n}{(n-k)k}$$

となるので，$f''(np)$ については，X の分散 $V(X)$ と標準偏差 $\sigma = \sigma(X)$ が $V(X) = \sigma^2 = npq$ であることに注意すると

$$f''(np) \fallingdotseq -\frac{n}{(n-np)np} = \frac{-1}{nqp} = \frac{-1}{\sigma^2}$$

が得られます．

よって，$f(k) = \log P(k)$ は $k = np$ 付近の 2 次近似で

$$f(k) \fallingdotseq f(np) - \frac{(k-np)^2}{2\sigma^2}$$

となり，したがって $f(np) = \log P(np)$ より

$$P(k) \fallingdotseq P(np)e^{-\frac{(k-np)^2}{2\sigma^2}} \ \left(= P(np)\exp\left[-\frac{(k-np)^2}{2\sigma^2}\right] \text{と書く}\right)$$

が得られます．記号 $\exp[x] = e^x$ は x が複雑な式のときに，指数関数を見やすくするためによく用いられます．

後は比例定数 $P(np)$ を求めれば完成です．そのために全確率が 1 であることを用いましょう：

$$1 = \sum_{k=0}^{n} P(k) \fallingdotseq \sum_{k=0}^{n} P(np)\exp\left[-\frac{(k-np)^2}{2\sigma^2}\right].$$

ここで，上式の k は整数なので，その増分 $\Delta k = (k+1) - k = 1$ を考えると

$$1 \fallingdotseq \sum_{k=0}^{n} P(np)\exp\left[-\frac{(k-np)^2}{2\sigma^2}\right]\Delta k$$

のように表すことができ，区分求積法のように短冊の面積の和の形に表すことができます．よって，k を実数に拡張すると，よい近似で積分の形

$$1 \fallingdotseq \int_0^n P(np)\exp\left[-\frac{(k-np)^2}{2\sigma^2}\right]dk$$

に直すことができます．ここで，置換

$$t = \frac{k - np}{\sqrt{2}\sigma} \qquad \begin{array}{c|ccc} k & 0 & \to & n \\ \hline t & t_0 & \to & t_n \end{array} \qquad \left(t_0 = \frac{-np}{\sqrt{2}\sigma}, \quad t_n = \frac{nq}{\sqrt{2}\sigma}\right)$$

を行うと

$$1 \fallingdotseq \int_{t_0}^{t_n} P(np) e^{-t^2} \sqrt{2}\sigma\, dt = \sqrt{2}\sigma P(np) \int_{t_0}^{t_n} e^{-t^2} dt$$

となります．ここで，$n \to \infty$ のとき $t_0 \to -\infty$，$t_n \to +\infty$ であり，また e^{-t^2} は，n が十分に大きいとき，$t < t_0$ および $t > t_n$ で無視できるほど小さくなります．よって，$P(np)$ を求めることは，$t_0 \fallingdotseq -\infty$ および $t_n \fallingdotseq \infty$ の近似を行って

$$\int_{t_0}^{t_n} e^{-t^2} dt \fallingdotseq \int_{-\infty}^{\infty} e^{-t^2} dt \ (= I \text{ とおく})$$

の計算に帰着します．

上の積分 I は君たちが大学で必ず演習する有名なもので

$$I = \int_{-\infty}^{\infty} e^{-t^2} dt = \sqrt{\pi}$$

となることが知られています．大学では華麗な方法で鮮やかに導きますが，ここでは高校生にも理解できる方法を用いましょう．e^{-t^2} は直接には積分できないので，何とか指数関数の形にもっていくための技巧を凝らします．

無限級数の積 $\sum\limits_k^{\infty} a_k \sum\limits_l^{\infty} b_l$ は，それらが絶対収束するとき，和をとる順序を変えても構いませんね：

$$\sum_k^{\infty} a_k \sum_l^{\infty} b_l = \sum_l^{\infty} \sum_k^{\infty} a_k b_l.$$

積分 $\int_a^b f(x)\,dx$ も区分求積の形 $\lim\limits_{\Delta x \to 0} \sum\limits_k f(x_k)\Delta x$ で考えると積分は無限級数になるので，積分の積を考えたときは積分を行う順序を変えることができます：

$$\int_a^b f(x)\,dx \int_c^d g(y)\,dy = \int_c^d \left\{ \int_a^b f(x) g(y)\,dx \right\} dy.$$

厳密な議論をしたい人は §§ 11.7.2 を参考にしてください．

e^{-t^2} は偶関数なので，$I = 2\int_0^\infty e^{-t^2}dt$ としておきましょう．まず，置換 $t = uv$（$v\,(>0)$ はパラメータ）を行います：

$$I = 2\int_0^\infty e^{-u^2v^2}v\,du.$$

次に，$I = 2\int_0^\infty e^{-v^2}dv$ とも表せることを利用して，I^2 を考えます：

$$I^2 = 4\int_0^\infty e^{-v^2}dv\int_0^\infty e^{-u^2v^2}v\,du.$$

そこで積分の順序を変えると

$$I^2 = 4\int_0^\infty\left\{\int_0^\infty e^{-(1+u^2)v^2}v\,dv\right\}du$$

となり，中括弧の中の積分：

$$J = \int_0^\infty e^{-(1+u^2)v^2}v\,dv$$

は，置換 $v^2 = s$ によって指数関数の積分の形になり，積分が実行できます：

$$J = \int_0^\infty e^{-(1+u^2)s}\frac{1}{2}ds = \left[-\frac{1}{2}\frac{e^{-(1+u^2)s}}{1+u^2}\right]_0^\infty = \frac{1}{2}\frac{1}{1+u^2}.$$

よって，

$$I^2 = 4\int_0^\infty \frac{1}{2}\frac{1}{1+u^2}\,du$$

となります．最後に，$u = \tan\theta$ と置換すると

$$\int_0^\infty \frac{1}{1+u^2}\,du = \frac{\pi}{2}$$

が得られます．これは練習問題にしましょう．したがって，

$$I^2 = 2\cdot\frac{\pi}{2} = \pi,\ \text{よって}\quad I = \int_{-\infty}^\infty e^{-t^2}dt = \sqrt{\pi}$$

が得られます．

以上のことから

$$1 \fallingdotseq \sqrt{2}\sigma P(np)\,I = \sqrt{2}\sigma P(np)\sqrt{\pi},\ \text{よって}\quad P(np) \fallingdotseq \frac{1}{\sqrt{2\pi}\,\sigma}$$

と決まり，最終的に二項分布 $P(X = k) = {}_n\mathrm{C}_k p^k q^{n-k}$ の n が非常に大きいときの近似の分布

$$P(X = k) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left[-\frac{(k - np)^2}{2\sigma^2} \right] \qquad (np = E(X), \quad \sigma = \sigma(X) = \sqrt{npq})$$

が得られます．この分布は **正規分布** (normal distribution) と呼ばれ, $N(np,\ \sigma^2)$ と表されます．np は X の平均値 $E(X)$，σ^2 は分散 $V(X)$ の値です．

なお, この近似を得るためにかなり荒っぽいことをやったように思えますが, 実際には $n \fallingdotseq 100$ になると極めてよい近似になり，実用上は $np > 5$，$nq > 5$ であれば，二項分布を正規分布として扱って差し支えないとされています．

15.5.3　正規分布

正規分布を応用するにあたって，その性質や特徴を調べておきましょう．

15.5.3.1　正規分布と中心極限定理

正規分布 $N(m,\ \sigma^2)$：

$$P(X = k) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left[-\frac{(k - m)^2}{2\sigma^2} \right] \qquad (m = E(X), \quad \sigma = \sqrt{V(X)})$$

を調べましょう．二項分布の近似のときは $m = np$，$\sigma = \sqrt{npq}$ $(q = 1 - p)$ です．正規分布では $P(X = k)$ の k は全実数に拡張され，全確率が 1 であることは，前の §§ の議論から確かめられるように，積分を用いて

$$\int_{-\infty}^{\infty} P(X = k)\,dk = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left[-\frac{(k - m)^2}{2\sigma^2} \right] dk = 1$$

のように表されます．二項分布のときは $0 \leqq k \leqq n$ でしたが，n が大きいときは $P(X < 0)$，$P(X > n)$ が非常に小さくなるので，$-\infty < k < \infty$ で議論したほうが都合がよいというわけです．

k が実数であることを表すには文字 x のほうが都合がよいので，以後正規分布 $N(m,\ \sigma^2)$ を表すときは

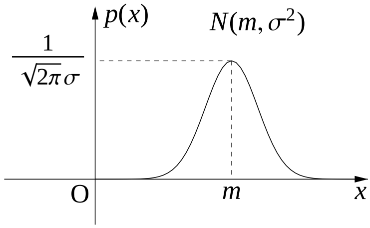
$$P(X = x) = p(x) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left[-\frac{(x - m)^2}{2\sigma^2} \right] \qquad (m = E(X), \quad \sigma = \sqrt{V(X)})$$

のようにしましょう．$p(x)$ はこの確率分布が連続分布の確率密度であることを表します．よって，$a \leqq X \leqq b$ である確率は積分を用いて

$$P(a \leqq X \leqq b) = \int_a^b p(x)\,dx = \int_a^b \frac{1}{\sqrt{2\pi}\,\sigma} \exp\Big[-\frac{(k-m)^2}{2\sigma^2}\Big]dx$$

と表されます[3)]．

右図の山型の分布から納得できるように，正規分布は，君たちの身長の分布や，君たちの成績の分布（ただし，極端に上位や下位のものを除く），工業製品の規格からのずれの分布であるとか，農産物などの大きさや重量の分布，種々の実験の測定値の分布など，多くの場合に適用できる最も重要な確率分布になっています．



さらに，大数の法則を精密化した「中心極限定理」という一般的な定理が導かれており，それによると n が十分に大きいときの分布はほとんどの場合に正規分布として扱うことができます．中心極限定理の証明には高度な数学を必要とするので，ここでは定理の内容をサンプル抽出の例や多くのサイコロを振る例で述べましょう．

n 個のサンプルのそれぞれの確率変数を X_1，X_2，$\cdots$，X_n として，それらは互いに独立でかつ共通の平均値 $E(X_i) = m$ と分散 $V(X_i) = \sigma^2$ をもつとします．各 X_i の確率分布 $P(X_i)$ は二項分布でなくても構いません．さて，大数の法則の場合と同じく，サンプルの平均をとる確率変数

$$\overline{X} = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

を導入すると，既に学んだように

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

3) $\sum_{k=0}^{n} P(k)$ を積分に直すときに，増分 $\varDelta k = 1$ を導入して，それを短冊の和 $\sum_{k=0}^{n} P(k)\varDelta k$ としたことを考えると，有限の n に対しては

$$P(a \leqq X \leqq b) = \int_{a-\frac{1}{2}}^{b+\frac{1}{2}} p(x)\,dx$$

とするほうがよりよい近似になるでしょう．

を得ます．中心極限定理は，n が十分に大きいとき，確率変数 $\overline{X}$ の確率分布 $P(\overline{X})$ が正規分布 $N(m, \frac{\sigma^2}{n})$ で近似されることを保証する定理です．また，サイコロを n 個振る例でいうと，個々のサイコロの目の数 X_i の分布は一様分布ですが，$\overline{X}$ を確率変数にすると分布は一様でなくなり，特に n が十分に大きくなると $\overline{X}$ の分布は正規分布に近づくというわけです．

この定理の重要な点は，確率変数として個々の X_i ではなく，それらの平均をとる確率変数 $\overline{X}$ を用いていることです．ただし，代わりに和をとる確率変数 $X = X_1 + X_2 + \cdots + X_n = n\overline{X}$ を用いても構いません．X を確率変数にすると，$E(X_i) = m$，$V(X_i) = \sigma^2$ のとき $E(X) = nm$，$V(X) = n\sigma^2$ となるので，中心極限定理は‘ n が十分に大きいとき X の確率分布は正規分布 $N(nm, n\sigma^2)$ に近づく ’という言い方になります．前の §§で行った二項分布の正規分布近似ではこの確率変数 X が用いられました．

15.5.3.2　正規分布の標準化

正規分布 $N(m, \sigma^2)$：

$$P(X = x) = \frac{1}{\sqrt{2\pi}\sigma}\exp\Big[-\frac{(x-m)^2}{2\sigma^2}\Big] \qquad (m = E(X), \quad \sigma = \sqrt{V(X)})$$

において，$a \leqq X \leqq b$ の確率を求めるためには

$$P(a \leqq X \leqq b) = \int_a^b \frac{1}{\sqrt{2\pi}\sigma}\exp\Big[-\frac{(x-m)^2}{2\sigma^2}\Big]dx$$

ですから，積分を行う必要があります．この積分の原始関数は求められないので，積分は区分求積などを用いる近似の積分になります．そのために確率変数

$$Z = \frac{X-m}{\sigma} \qquad \Big(z = \frac{x-m}{\sigma}\Big)$$

を用いましょう．すると，確率 $P(a \leqq X \leqq b)$ は

$$z_a = \frac{a-m}{\sigma}, \qquad z_b = \frac{b-m}{\sigma}$$

とおくと，$z_a \leqq Z \leqq z_b$ に対する確率

$$P(z_a \leqq Z \leqq z_b) = \int_{z_a}^{z_b} \frac{1}{\sqrt{2\pi}}\exp\Big[-\frac{z^2}{2}\Big]dz$$

として表されます．よって，Z の確率密度 $P(Z=z)=p(z)$ は

$$p(z)=\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{z^2}{2}\right]$$

となります．この確率分布は

$$E(Z)=\frac{E(X)-m}{\sigma}=0,\qquad V(Z)=\frac{1}{\sigma^2}V(X-m)=1$$

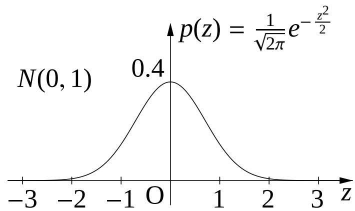
となるので，**標準正規分布** $N(0,1)$ といわれます．

確率 $P(z_a\leqq Z\leqq z_b)$ は，確率

$$P(Z\leqq z)=\Phi(z)=\int_{-\infty}^{z}\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{z'^2}{2}\right]dz'$$

の表を近似積分によって作っておくと

$$P(z_a\leqq Z\leqq z_b)=\Phi(z_b)-\Phi(z_a)$$



から求めることができます．$\Phi(z)$ の表はかなり膨大であり，また君たちの教科書の巻末付録に載っていますので，ここでは割愛しますが，それから得られる重要な確率の例を挙げておきましょう（以下，$\Phi(z)$ の数値を用いたところでは，= は ≒ を意味します）：

$$P(|X-m|\leqq 1\sigma)=P(|Z|\leqq 1)=0.683,$$

$$P(|X-m|\leqq 2\sigma)=P(|Z|\leqq 2)=0.954,$$

$$P(|X-m|\leqq 3\sigma)=P(|Z|\leqq 3)=0.997.$$

これらの数値からわかることは，$|X-m|\leqq 1\sigma$ である確率が既に 68 %を占め，$|X-m|\leqq 3\sigma$ では 99.7 %と実質的に 100 %となることです．また，サンプル抽出の際の信頼度を調べるのによく用いられるのは確率が 95 %となる場合で，厳しく 99 %とする場合もあります：

$$P(|X-m|\leqq 1.96\sigma)=P(|Z|\leqq 1.96)=0.95,$$

$$P(X-m\leqq 1.65\sigma)=P(Z\leqq 1.65)=0.95,$$

$$P(|X-m|\leqq 2.58\sigma)=P(|Z|\leqq 2.58)=0.99.$$

ここで，正規分布の簡単な応用例として大学入試センター試験のときなどで用いられる「偏差値」を議論しておきましょう．試験の点数は 100 点満点とは限りませんし，平均点も 50 点となることはまずありませんね．偏差値は平均が 50，標準偏差が 10 となるように点数を換算した値です．

君たちの点数を X として，X の分布は，平均値 m，標準偏差 σ の正規分布になっているとしましょう．このとき，確率変数

$$Y = 50 + \frac{X - m}{\sigma} \times 10 = 50 + Z \times 10$$

を導入すると，

$$E(Y) = 50 + \frac{E(X) - m}{\sigma} \times 10 = 50,$$

$$V(Y) = \frac{10^2}{\sigma^2} V(X - m) = 10^2,\ \text{よって} \quad \sigma(Y) = 10$$

となって，点数がうまく換算されます．

そこで，君の偏差値が 65 であったとすると，$\Phi(z)$ の表から

$$P(Y \leqq 65) = P(Z \leqq 1.5) = \Phi(1.5) = 0.9332,$$

よって，$P(Y \geqq 65) = 0.0668$ となります．これは，君が上位から 6.68 %の順位にいることを表します．もし千人の受験希望者がいるとすると，君は上位から $1000 \times 0.0668 \fallingdotseq 67$ 番目であることがわかります．もし偏差値が 70 なら，$Z \leqq 2$ なので $\Phi(2) = 0.9772$ を用い，上位から $(1 - 0.9772) \times 100 = 2.28$ %の順位となります．

ただし，偏差値に一喜一憂する必要はありません．努力次第で学力は伸びていきます．‘センター試験で計ることができる学力’は‘高校レベルで必要とされる学力に対してどれほど要領よく学んでいるか’であって，君が‘持続的努力によって獲得できるであろう真の能力を計れるような代物では決してないのです’．うまく第 1 希望の難関大学に入ったとしても，それで君の人生が死ぬまで巧(うま)くいくと保証されているなどと考えるのはまったくの幻想です．君が社会人になったら，何ヵ月も何年もかけてようやく解決できるような一連の問題に遭遇するでしょう．もっとも，そのような問題は，他人から与えられるというよりは，自分で作り出すようなものでしょうが．目先の試験にあまり囚わ

れずに君の真の能力を開発するための努力を続けましょう．そのキーワードは“ 自分は，何のために生まれ，何のために生きているのか ”です．

15.5.4　視聴率

15.5.4.1　視聴率の推定

既に触れた視聴率の問題を詳しく調べてみましょう．人気テレビ番組 A の真の視聴率を（未知の値）p として，これをテレビをもつ全世帯から n 世帯をサンプルとして抽出します．番組 A を見ていた世帯数を X とすると，$X = k$ である確率は二項分布 $B(n,\ p)$：

$$P(X = k) = {}_n\mathrm{C}_k p^k (1-p)^{n-k}, \qquad (E(X) = np, \quad \sigma(X) = \sqrt{npq})$$

で表され，サンプル数 n が十分に大きいとき，それは正規分布 $N(np, \sigma^2)$：

$$P(X = k) = \frac{1}{\sqrt{2\pi}\sigma} \exp\Big[-\frac{(k-np)^2}{2\sigma^2}\Big] \qquad (\sigma^2 = np(1-p))$$

で非常によく近似されますね．

番組 A を見ていた世帯数がある調査で $X = k_s$ であるとすると，n が十分に大きいときは，大数の法則により，よい近似で $k_s \fallingdotseq E(X) = np$ が成り立つと考えられます．よって視聴率 p と分散 σ^2 が

$$p \fallingdotseq \frac{k_s}{n}, \qquad \sigma^2 \fallingdotseq \frac{k_s(n-k_s)}{n} \ \Big(= \sigma_s^2 \text{ とおく}\Big)$$

と推定されます．このような推定を **点推定** といいます．マスメディアが発表する視聴率はこの点推定による視聴率です．

点推定視聴率 $\frac{k_s}{n}$ がどれほど信頼できるかを確かめるにはサンプルを何度もとり直せばよいのですが，それは事実上不可能でしょう．それに代わるものとして確率分布と視聴世帯数 $X = k_s$，点推定による標準偏差 $\sigma \fallingdotseq \sigma_s$ を利用する方法があります．例えば，前の §§の議論より，95 %確かである確率は，$\sigma \fallingdotseq \sigma_s$ より

$$P(|X - np| \leqq 1.96\sigma_s) = 0.95$$

となる X の範囲となり，これを利用すると 95 %の信頼度で真の視聴率 p の範囲がわかります：

$$\begin{aligned}|X-np| \leqq 1.96\sigma_s \ &\Leftrightarrow \ -1.96\sigma_s \leqq X-np \leqq 1.96\sigma_s \\ &\Leftrightarrow \ X-1.96\sigma_s \leqq np \leqq X+1.96\sigma_s\,.\end{aligned}$$

ここで，調査で得られた値 $X=k_s$ は確率が 95 %である X の範囲にあると考えられるので，不等式

$$k_s-1.96\sigma_s \leqq np \leqq k_s+1.96\sigma_s \qquad \text{（区間推定）}$$

が 95 %の確かさで成り立ちます．この不等式を n で割ると p が 95 %の信頼度で推定できます．このように，ある範囲を与えるような推定を **区間推定** といいます．

ある代表的な調査会社はサンプル数 n が関東・関西地区で 600 世帯とのことです．ある時間の視聴世帯数 X が $k_s=200$ とすると，点推定視聴率は 33 %となります．これを 95 %の信頼度で区間推定してみましょう．上の（区間推定）の式で，$n=600$，$k_s=200$ を代入すると，

$$\sigma_s = \sqrt{\frac{k_s(n-k_s)}{n}} = 11.547$$

なので，

$$0.30 \leqq p \leqq 0.37$$

が得られ，真の視聴率 p は 95 %の信頼度で 30 %から 37 %の間にある，つまり $p=33.5\pm3.5$ % と誤差が僅かに 3.5 %であることがわかります．

では，ここで練習をします．上の問題を 99 %の信頼度で区間推定せよ．ヒント：上の（区間推定）の式の 1.96 を 2.58 で置き換えればよいのですが，意味を理解するために前の §§ から読み返しましょう．$P(|X-np| \leqq 2.58\sigma)=0.99$ より同様にして

$$k_s-2.58\sigma_s \leqq np \leqq k_s+2.58\sigma_s$$

となるので，$n=600$，$k_s=200$，$\sigma_s=11.547$ を代入して $0.28 \leqq p \leqq 0.38$ が得られるので，$p=33\pm5$ % となりますね．

15.5.4.2　視聴率の検定

さて，あるハンサムな人気タレントのドラマはいつも高視聴率を維持し，今週は最高視聴率の 40 %を記録しました．次の日，40 人学級のあるクラスでそのドラマを見た人を調べたところ 20 人いて，そのクラスの視聴率は 50 %でした．ところが隣のクラスで調べてみると 12 人しかおらず，そのクラスでは 30 %です．真の視聴率 p が 40 %のとき，サンプル数が 600 のときは 95 %の信頼度で視聴率の誤差は 2 %程度です．40 人学級で視聴率が 50 %とか 30 %という結果になることはあるのでしょうか．調べてみましょう．

40 人学級のあるクラスを選ぶのは，視聴率が p である母集団からランダムに 40 世帯を抽出したと考えられますね．$n = 40$ 人のうち $X = k$ 人が見た確率分布は，彼らが個人の好みで見たとすると，二項分布 $B(n,\ p)$ で表され，また $n = 40$ は大きいのでそれは正規分布 $N(np,\ \sigma^2)$ で近似されますね．

さて，真の視聴率 p が 40 %のとき，40 人学級で視聴率が 50 %とか 30 %という結果，つまり 40 人中 20 人が見たとか 12 人が見たとかいうのは，あり得ないとはいいませんが‘ 妥当ではない ’と考えてみましょう．このことは $X = 20$ や $X = 12$ が平均値 $E(X) = np$ の近くにはないことを意味します．確率でいうと，$N(np,\ \sigma^2)$ $(n = 40,\ p = 0.4,\ \sigma^2 = np(1-p))$ において，$X = 20$ や $X = 12$ を含む区間 $|X - np| \geqq z\sigma$ の確率の値が小さい：

$$P(|X - np| \geqq z\sigma) = \alpha\ (\fallingdotseq 0)$$

と表すことができます．

この α の値をいくらにとればよいかは重要で，その数値は意義がある標準と考えられる値の意味で **有意水準** といわれます．多くの場合に理に叶う値と考えられるのは有意水準 $\alpha = 0.05$ で，40 人学級の視聴率が 50 %以上または 30 %以下となる確率は 5 %もないことを表します．$\alpha = 0.05$ とすると，前の§§の議論から

$$P(|X - np| \geqq 1.96\sigma) = 0.05$$

が成り立つので，$X = 20$ や $X = 12$ は区間 $|X - np| \geqq 1.96\sigma$ にあるはずです．

この区間は，$n = 40$，$p = 0.4$ ですから，$\sigma = \sqrt{9.6} = 3.1$ より

$$|X - 16| \geqq 6.1$$

となり，$X = 20$ や $X = 12$ を代入してみると，共に $4 \geqq 6.1$ となって不等式を満たしません．この結果は，40 人学級の視聴率が 50 %とか 30 %となるのは妥当でないとする‘ 仮説 ’は誤りであり，その仮説は‘ 捨てるべき ’であると判定されたことを意味します．このような判定の仕方を仮説の **検定** といいます．

以上の議論から，$X = 20$ や $X = 12$ は

$$P(|X - np| \leqq 1.96\sigma) = 0.95$$

が成り立つ平均値 $E(X) = np$ の近くの区間 $|X - np| \leqq 1.96\sigma$ にあることになります．このことを，先に行った区間推定の議論に従って調べてみましょう．$n = 40$，$X = k_s = 20$ とすると，$p \fallingdotseq 0.5$ として $\sigma_s = 3.16$ となりますから，上の不等式は，区間推定法より

$$|20 - 40p| \leqq 6.2$$

と表され，これを解いて，視聴率が 95 %の信頼度で

$$0.35 \leqq p \leqq 0.65$$

と推定され，真の視聴率 $p = 40$ %を含みます．この結果は $p = 50 \pm 15$ %と誤差が大きく，$n = 40$ 程度のサンプル数ではばらつきが大きくてあまり信用できる結果は得られないことを意味します．つまりこのクラスの視聴率が 50 %となったことは驚くに当たらないということです．同様のことは $X = 12$ としてもいえますが，それを確かめるのは君たちに任せましょう．

なお，上の問題で，視聴率が $p = 40$ %であるとき，ある 40 人学級のクラスで 20 人もの人が見たというのは妥当であるか，と問題を設定すると検定の仕方が変わります．この問題設定では見た人数が多いことを問題にし，少ないのは構わないというわけです．よって，有意水準 5 %の検定を行うには，今度は

$$P(X - np \leqq 1.65\sigma) = 0.95$$

から得られる

$$P(X - np \geqq 1.65\sigma) = 0.05$$

を用います．先の議論と同様にして，$n = 40$，$X = 20$，$p = 0.4$，$\sigma = 3.1$ より，不等式

$$20 - 16 \geqq 5.1$$

が得られます．これは満たされないので，20 人が見たというのは多すぎるという仮説は捨てることになります．このような検定を「片側検定」といい，先に行った平均値からのずれを調べる検定を「両側検定」といいます．