CG のための確率

mebiusbox

2022 年 12 月 28 日

はじめに

プログラミングにおいて，処理の分岐や度合いを確率に基づいて決定することはよくあります．例えば，スマホゲームでよくあるガチャの提供割合や， AI の行動を確率によって分岐していることもあります．また，迷惑メールの判定でも確率は使われています．そして，CGでもレンダリングの計算時に確率が使われることがあります．そこで，確率を扱うときに重要な確率分布について基礎から理解できるようにまとめてみました．確率分布はモンテカルロ法によるレイトレーシングのサンプリングや，物理ベースレンダリングのライティング計算における反射の割合，分散シャドウマップ法による遮蔽計算などで出てきます．確率分布にはその特徴を表す平均値や分散というのがあり，それらを調整したり，分布そのものを変えることで結果を変えるといったことができます．

今回は，確率の基礎である場合の数からはじめ，確率分布，平均値，分散，そしていくつかの有名な確率分布について触れていきます．

目次

1. 確率の考え方

世の中に明日(未来)のことが分かる人がいるなんて思いませんし,私や皆さんも分からないと思います.明日が平日の場合,社会人なら会社に出社し,学生なら学校に行くでしょう.もしかすると休日の人もいるかもしれません.また,何かしらの原因で交通手段がなくなってしまい,駅で待つか,自宅待機になるかもしれません.明日の予定は決まっているけど,明日になってみないと予定どおりにいくかどうかはわからないわけです.

ここで,社会人に対して「あなたは明日通常どおり出社できますか?」と聞かれたらどう答えるでしょうか.まあ,ほとんどの人は「できる」と答えるでしょうね.そこで,次のような状況を考えます.あなたは,電車を使って出勤しています.最寄りの駅を使いますが,この駅から発車する電車は一日1本です.極端ですね.しかも,この電車は数日に1回は故障してその日は走行できないポンコツです.そんなところに住まないし,電車を使わずに車や自転車を使うって野暮な選択はなしです.あと,寝坊とかしないでその電車の発車に間違いなく間に合うマジメ君とします.その状況で,もう一度「あなたは明日通常どおり出社できますか?」と聞かれたらどうでしょうか.まあ,私だったら「わからない」でしょうね.では,数日に1回ではなく100日に1回故障するということでしたらどうでしょう.「多分大丈夫」と答えるかもしれません.さらに,故障が100年に1回なら「出社できる」と答えそうですね.

ここで注目したいのがある期間中に故障する回数がわかっていると,明日故障するかどうかの予測が直感的にできるということです.もちろん,実際に故障するかどうかは偶然によって決まりますし,予測がはずれることもあります.ですが,ある程度予測できるなら対応もやりやすくなります.そのための道具が確率です.

もし,明日の故障が高い確率で起きるということがわかっていれば,人によっては電車ではなく車や自転車を使って出社しようと考えるかもしれません.また,高い確率という表現は結構曖昧で,その確率を定量的に表現できれば,もっと的確になります.例えば,その電車について 100 日に 1 回故障したデータがあった場合,明日は 1/100 の確率で故障すると言えます.これは

$$\text{故障する確率} = \frac{\text{故障した回数}}{\text{観測した日数}}$$

と計算した値です.これを一般的にいえば,ある事象が発生する確率は

$$\text{確率} = \frac{\text{ある事象が起こる場合の数}}{\text{すべての事象が起こる場合の数}}$$

となります.先程の例でいうと,ある事象が起こる場合の数は故障した回数で,すべての事象が起こる場合の数は観測した日数ということになります.

このように確率を計算するためには,ある事象の場合の数とすべての事象の場合の数が必要です.この2つの場合の数を正確に数え上げることが大切です.

2. 集合

2.1 集合と要素

互いに識別できる「もの」の集まりを,集合といいます.たとえば,1から5までの自然数 1, 2, 3, 4, 5 の集まりは集合です.この場合,$\{1, 2, 3, 4, 5\}$,あるいは $A = \{1, 2, 3, 4, 5\}$ と表します.1つひとつを要素といいます.要素をまったくもたない集合のことを空集合といい,ϕ で表されます.また,ある集合 A の要素の個数は $n(A)$ と表します.

2.2 共通部分と和集合

集合 B の要素がすべて集合 A の要素であるとき,B は A の部分集合であるといいます.そして $B \subset A$,または $A \supset B$ と表します.

一般に, 集合 U を全体の集合とし, この中で部分集合を考えるとき, U を全体集合といいます. 部分集合 A と B の要素を合わせた集合を和集合といい, $A \cup B$ と表します. また, 部分集合 A と B のどちらにも入っている共通の要素を共通要素といい, $A \cap B$ と表します. 全体集合 U のうち, 部分集合 A ではない要素を A の補集合といい, $\overline{A}$ と表します.
全体集合 U, 部分集合 A, 補集合 $\overline{A}$ は次のような関係になります.

$$A \cup \overline{A} = U, \quad A \cap \overline{A} = \phi$$

3. 場合の数

確率の計算において, 場合の数を正確に数え上げることが大切です. 実際に場合の数を数え上げる場合には, 複雑で面倒な条件がついた場合の数が何通りで, その場合の数を考えたりして数え上げなければなりません. つまり, いかに場合の数を求めるかが確率計算の基本となります. まずは基礎となる2つの法則を見ていきます.

3.1 和の法則

例題 大小2つのサイコロを同時に振るとき, 出る目の数の和が 4 または 5 になる場合の数

サイコロの出る目は $\{1, 2, 3, 4, 5, 6\}$ です. 大小2つのサイコロの目の和が4になる出目は

大	1	2	3
小	3	2	1

大小2つのサイコロの目の和が5になる出目は

大	1	2	3	4
小	4	3	2	1

目の和が4になる事象を A, 目の和が5になる事象を B とします. 事象 A が起こる場合は3通りです. 場合の数を $n(A)$ で表すと, $n(A) = 3$ となります. 同じように事象 B が起こる場合は4通りなので $n(B) = 4$ です. ここで, 事象 A と事象 B は同時に起こらないことがわかります. このような関係の事象は互いに排反であるといいます.
求めたい事象は「目の和が4か5になる事象」なので, この2つの事象が同時に起こらないならば, この事象が起こる場合の数は

$$n(A) + n(B) = 3 + 4 = 7$$

となります. これを和の法則といいます. これは3つ以上の事象でも成り立ち, 一般に事象 $A, B, \ldots, N$ があり, どの2つの事象も同時に起こらないならば, 事象 $A, B, \ldots, N$ のどれかが起こる場合の数は

$$n(A) + n(B) + \cdots + n(N)$$

となります.

3.2 積の法則

例題 伊藤, 加藤, 佐藤の3種類の姓と, 太郎, 花子の2種類の名前を組み合わせて, 何通りの人名が作れるか

3種類の姓の事象を A, 2種類の名を事象 B とします.姓から1つを選ぶと, それに対して2つの名を選ぶことができます.他の2つの姓でもそれぞれに2つを選ぶことができます.結局, 次のように人名が6通り作れます.

	太郎	花子
伊藤	伊藤太郎	伊藤花子
加藤	加藤太郎	加藤花子
佐藤	佐藤太郎	佐藤花子

これは, 事象 A の起こり方が $n(A)$ 通りあり, その各々に対して事象 B の起こり方が $n(B)$ 通りあるなら, 事象 A と B がともに起こる場合の数は

$$n(A) \times n(B) = 3 \times 2 = 6$$

になります.これを積の法則といいます.これも3つ以上の事象でも成り立ち, 一般に事象 $A, B, \ldots, N$ があり, 事象 $A, B, \ldots, N$ がともに起こる場合の数は

$$n(A) \times n(B) \times \cdots \times n(N)$$

となります.

例題 1, 2, 3, 4 の4個数字を使ってできる整数で, 各けたの数字が異なる4けたの整数は何個ありますか

1度使った数字は別のけたで使用できないことを考慮して, まず千の位で使える数字は4個全部なので4通り, 次に百の位では千の位で使った数字以外の中から選ぶので3通り, 同じように十の位では千の位と百の位で使った数字以外なので2通り, 最後に一の位では残りの数字なので1通りとなります.表で表すと次のようになります.

千の位	百の位	十の位	一の位
4通り	3通り	2通り	1通り

千の位, 百の位, 十の位, 一の位での場合の数をそれぞれ事象 A, B, C, D として積の法則を使うと, 各けたの数字が異なる4けたの整数の数は

$$n(A) \times n(B) \times n(C) \times n(D) = 4 \times 3 \times 2 \times 1 = 24$$

となります.

4. 順列

ある集合について, それらを順序をつけて並べることを考えます.例えば {雪, 月, 花} という集合があったとき,

雪, 月, 花　雪, 花, 月

月, 雪, 花　月, 花, 雪

花, 月, 雪　花, 雪, 月

の6通りあります.一般にいくつかのものを順序をつけて1列に並べた配列を**順列**といいます.

ここで,積の法則で出てきた例題をもう一度見てみます.

例題 $1, 2, 3, 4$ の4個数字を使ってできる整数で,各けたの数字が異なる4けたの整数は何個ありますか

この整数を列挙してみると次のようになります.

4321	4312	4231	4213	4123	4132
3421	3412	3241	3214	3124	3142
2341	2342	2431	2413	2143	2134
1324	1314	1234	1243	1423	1432

これは4つの数字の順列になっています.ここで4個ではなく n 個の数字を使って,各けたの数字が異なる n けたの整数の数は

$$n \times (n-1) \times \cdots \times 3 \times 2 \times 1$$

で求まります.この式は 1 から n までのすべての自然数の積で $n!$ で表し,これを**階乗**といいます.

$$n! = 1 \times 2 \times 3 \times \cdots \times (n-1) \times n$$

例題の場合, $n = 4$ なので

$$4! = 4 \times 3 \times 2 \times 1 = 24$$

と表します.

例題 $\{1, 2, 2, 2, 3, 4\}$ の数字を左から順に1列に並べてできる整数は何個ありますか

今度は6桁の整数で,同じ数字が含まれています.2が3つありますが,これら3つを互いに区別できるものとして考えると,先程の例題と同じように個数を求められるので $n = 6$ として

$$6! = 6 \times 5 \times 4 \times 3 \times 2 \times 1 = 720$$

となります.例えば 321224 という数字を選んだとき,「3 ● 1 ●● 4」の●をすべて区別して数えていることになります.この場合,●の順列 $3!$ 通りを1通りとして数え直す必要があります.結局,求める場合の数は

$$\frac{6!}{3!} = \frac{720}{6} = 120$$

となります.

この同じものを複数個含んだ順列を一般式で考えてみます.

例題 n 個のもののうちで, m_1 個, m_2 個, $\ldots$, m_k 個がそれぞれ同じ種類のものであるとき,この n 個で作られる順列の総数はいくつになりますか.ただし $n = m_1 + m_2 + \cdots + m_k$ である.

求めたい場合の数を x とします.ここで, m_1 個が互いに区別できるものとして考えて並べると $m_1!$ 通りあることになります. $m_2, \ldots, m_k$ も $m_2!, \ldots, m_k!$ とし,積の法則を使って次のように計算します.

$$x \times m_1! \times m_2! \times \cdots \times m_k!$$

この数は n 個のすべてが異なっているものとした順列の総数と一致します.よって,

$$x \times m_1! \times m_2! \times \cdots \times m_k! = n!$$

ここで, 式を変形して x を求めると

$$x = \frac{n!}{m_1! \times m_2! \times \cdots \times m_k!}$$

となります.

例題 MISSISSIPPI というアルファベットを並び替えてできる単語の総数はいくつになりますか

アルファベットの総数は 11 個で, I が4つ, S が4つ, P が2つなので, $n = 11, m_1 = 4, m_2 = 4, m_3 = 2$ となって

$$\frac{11!}{4!4!2!} = \frac{11 \times 10 \times 9 \times 8 \times 7 \times 6 \times 5}{4 \times 3 \times 2 \times 2} = 34650$$

となります.

例題 1から6の数字が書かれたカードが6枚あります.この6枚のカードから3枚選んで並べる方法は何通りありますか

最初に1枚選ぶときは6通り, 次に選ぶときは5通り, 3回目は4通りとなるので

$$6 \times 5 \times 4 = 120$$

もし3枚ではなく4枚選んで並べる方法の場合

$$6 \times 5 \times 4 \times 3 = 360$$

となります.これは n 個の中から r 個のものを選ぶ場合

$$n(n-1)(n-2) \cdots (n-r+1)$$

と表せます.ここで, 先程の $6 \times 5 \times 4$ は

$$\frac{6 \times 5 \times 4 \times 3 \times 2 \times 1}{3 \times 2 \times 1}$$

と変形することができます.階乗を使えば

$$6 \times 5 \times 4 = \frac{6!}{3!} = \frac{6!}{(6-3)!}$$

一般式にまとめると

$$n(n-1)(n-2) \cdots (n-r+1) = \frac{n!}{(n-r)!}$$

となります.順列の数は, 順列(Permutation)の頭文字 P を使って

$$\frac{n!}{(n-r)!} = {}_nP_r$$

と表記します.ここで, $r=0$ のとき

$${}_nP_0 = \frac{n!}{(n-0)!} = \frac{n!}{n!} = 1$$

となります.また, $r=n$ のとき

$${}_nP_n = \frac{n!}{(n-n)!} = \frac{n!}{0!}$$

となりますが, これは n 個のものから n 個を選んで並べる方法なので, n 個の順列の総数と同じです.よって

$${}_nP_n = n!$$

となります.これにより $0! = 1$ でなければなりません.

例題 1から9までの数字から異なる3個の数字を選んで3桁の整数をつくるとき, その総数はいくつになりますか

$n=9, r=3$ なので

$${}_9P_3 = 9\times 8\times 7 = 504$$

例題 A から Z までの 26 文字を使って 4 文字からなる単語を作るとき, 何通りの単語が作れますか.同じアルファベットを複数使ってもいい.

4 文字のうち, 最初の文字は 26 通り, 次の文字も 26 通り, 他の2文字も 26 通り選べます.積の法則から総数は

$$26\times 26\times 26\times 26 = 26^4$$

一般に n 個の異なるものから重複を許して r 個を取り出して並べてできる順列の総数は n^r で, ${}_n\Pi_r$ と表します.

$${}_n\Pi_r = n^r$$

5. 組み合わせ

順列では並び順も考えて場合の数を計算しましたが, 並び順は考えないで場合の数を計算することもあります.

例えば, 3枚のカードから3枚を選ぶ組み合わせは1通りです.これが順列なら $3! = 6$ 通りとなります.

それでは, この組み合わせを求める一般式を考えてみます.異なる n 個のものから r 個を取り出した組み合わせの数を組合せ(Combination) の頭文字 C を使って ${}_nC_r$ と表します.組合せの数に $r!$ を掛ければ, 積の法則より

$${}_nC_r \times r!$$

となって, これは全体の順列の数と一致します.よって

$$ {}_nC_r \times r! = {}_nP_r $$

式を変形して組合せの数を求めると

$$ {}_nC_r = \frac{{}_nP_r}{r!} = \frac{n!}{r!(n-r)!} $$

となります.ここで, r に $n-r$ を代入すると

$$ {}_nC_{n-r} = \frac{n!}{(n-r)!(n-(n-r))!} = \frac{n!}{(n-r)!r!} = {}_nC_r $$

また, $n=n$ とすると

$$ {}_nC_n = \frac{n!}{n!(n-n)!} = \frac{n!}{n!} = 1 $$

例題 10 枚のカードから 4 枚のカードを取り出したときの組合せは何通りありますか

$n=10, r=4$ なので,

$$ {}_{10}C_4 = \frac{10\times 9\times 8\times 7}{4\times 3\times 2} = \frac{5040}{24} = 210 $$

例題 1 から 9 の数字から 8 個の数字を取り出したときの組合せは何通りありますか

$n=9, r=8,$ また, ${}_nC_r = {}_nC_{n-r}$ を使うと

$$ {}_9C_8 = {}_9C_{9-8} = {}_9C_1 = \frac{{}_9P_1}{1!} = 9 $$

例題 9人を3組のグループに分ける方法は何通りありますか.ただし1組の人数は2人以上とする

3つの組の人数をそれぞれ l, m, n とし, それらの組の組合せの数を (l, m, n) として表すと, $(5,2,2), (4,3,2), (3,3,3)$ の3通りの人数の組合せがあります. $(5,2,2)$ の場合, まず 9 人から 5 人, 次に 4 人から 2 人, 残り 2 人の組合せの総数をそれぞれ計算し, 積の法則で総数を計算します.次に, 2人の組が2つありますが, 入れ替わっても組合せは同じなので, $2!$ 通り余計に数え上げられていることになります.そのため, 総数を $2!$ で割って, 調整します.

$$ (5,2,2) = \frac{{}_9C_5 \times {}_4C_2 \times {}_2C_2}{2!} = \frac{126\times 6\times 1}{2} = 378 $$

同様に $(4,3,2), (3,3,3)$ の場合を計算します.

$$ (4,3,2) = {}_9C_4 \times {}_5C_3 \times {}_2C_2 = 126\times 10\times 1 = 1260 $$

$$ (3,3,3) = \frac{{}_9C_3 \times {}_6C_3 \times {}_3C_3}{3!} = \frac{84\times 20\times 1}{6} = 280 $$

最後に和の法則から, 3組の場合の数を足し合わせて

$$ (5,2,2) + (4,3,2) + (3,3,3) = 378 + 1260 + 280 = 1918 $$

となります.

例題 赤, 青, 緑の3つの色がついた玉があります.この中から4つの玉を取り出したときの組合せは何通りあるでしょうか.同じ色の玉は何度でも取り出せることとします

すべての組合せを列挙してみると

赤赤赤赤	赤赤赤青	赤赤赤緑
赤赤青青	赤赤青緑	赤赤緑緑
赤青青青	赤青青緑	赤緑緑緑
赤緑緑緑	青青青青	青青青緑
青青緑緑	青緑緑緑	緑緑緑緑

このように 15 通りあります.これを一般式にしてみましょう. n 個の異なるものから, 重複を許して r 個取り出す組合せを考えます. n 個のものを $n-1$ の点(·)で区切ってみます.

赤赤赤赤··	赤赤赤·青·	赤赤赤··緑
赤赤·青青·	赤赤·青·緑	赤赤··緑緑
赤·青青青·	赤·青青·緑	赤·青·緑緑
赤··緑緑緑	·青青青青·	·青青青·緑
·青青·緑緑	·青·緑緑緑	··緑緑緑緑

次に r 個の取り出したものを●に置き換えます.

●●●●··	●●●·●·	●●●··●
●●·●●·	●●·●·●	●●··●●
●·●●●·	●·●●·●	●·●·●●
●··●●●	·●●●●·	·●●●·●
·●●·●●	·●·●●●	··●●●●

この●と·の並び方はすべて異なっています.この並べ方は組合せと1対1に対応するから, $n-1$ 個の点と r 個の●の並べ方の総数を求めればいいことなります.これは $n-1$ 個の点と r 個の●を含めた全部のところから r 箇所の場所を取り出す組合せと考えることができます.各箇所に番号を割り振ると ①②③④⑤⑥ となって

①②③④	①②③⑤	①②③⑥
①②④⑤	①②④⑥	①②⑤⑥
①③④⑤	①③④⑥	①③⑤⑥
①④⑤⑥	②③④⑤	②③④⑥
②③⑤⑥	②④⑤⑥	③④⑤⑥

と表せます. n 個の異なるものから, 重複を許して r 個とる組合せを ${}_nH_r$ で表し

$$ {}_nH_r = {}_{n+r-1}C_r $$

となります.

例題 赤, 青, 緑の3つの箱に, 区別のつかない7個の玉を入れる場合, 何通りありますか.また, どの箱にも少なくとも1個の玉を入れるとすれば, 何通りの入れ方がありますか

重複を許した組合せなので, $n = 3, r = 7$ として

$$ {}_3H_7 = {}_{3+7-1}C_7 = {}_9C_7 = {}_9C_2 = \frac{9 \times 8}{2 \times 1} = 36 $$

となります.また, どの箱に少なくとも1個の玉を入れる場合は, あらかじめ各箱に1個の玉を入れておき, 残り4個の玉を3つの箱に入れることを考えて

$$ {}_3H_4 = {}_{3+4-1}C_4 = {}_6C_4 = {}_6C_2 = \frac{6 \times 5}{2 \times 1} = 15 $$

となります.

例題 異なる3つの箱に玉を分けるとき, 赤い玉が5個の場合と赤い玉が5個と白い玉が2個の場合の分け方は何通りありますか.玉を入れない箱があってもいいとします.

赤い玉が5個の場合は ${}_3H_5$ なので

$$ {}_3H_5 = {}_{3+5-1}C_5 = {}_7C_5 = {}_7C_2 = \frac{7 \times 6}{2} = 21 $$

となります.次に, 赤い玉が5個と白い玉が2個の場合は, まず白い玉が2個を異なる3つの箱に入れる場合の数は ${}_3H_2$ なので

$$ {}_3H_2 = {}_{3+2-1}C_2 = {}_4C_2 = \frac{4 \times 3}{2} = 6 $$

赤い玉が5個の場合と白い玉が2個の場合の積となるので $21 \times 6 = 126$ となります.

6. 2項定理

展開公式としてよく知られたものに

$$ \begin{aligned} (a+b)^2 &= a^2 + 2ab + b^2 \\ (a+b)^3 &= a^3 + 3a^2b + 3ab^2 + b^3 \\ (a+b)^4 &= a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + b^4 \end{aligned} $$

があります.これを一般化して, $(a+b)^n$ の展開を示す公式を2項定理といいます.

$n = 1, 2, 3, \ldots$ のとき

$$ (a+b)^n = (a+b)(a+b)(a+b)\cdots(a+b) $$

となりますが, これを展開すると, 各項は $a^{n-r}b^r(r = 0, 1, 2, \dots, n)$ の形になります. $a^{n-r}b^r$ の係数は, n 個の因数 $(a + b)$ から, b を r 個選ぶ組合せの数 ${}_nC_r$ となります.したがって, $a^{n-r}b^r$ の係数は ${}_nC_r$ となり, これを 2 項係数ともいいます. 2 項係数は $\binom{n}{r}$ と表記することもあります.

2 項定理をまとめると次のようになります.

$$
\begin{aligned}
(a+b)^n &= {}_nC_0a^n + {}_nC_1a^{n-1}b + {}_nC_2a^{n-2}b^2 + \cdots + {}_nC_ra^{n-r}b^r + \cdots + {}_nC_nb^n \\
&= \sum_{r=0}^{n} {}_nC_ra^{n-r}b^r
\end{aligned}
$$

少し寄り道して, この展開式を関数のべき級数展開を使って導出してみます.まずは $(1 + x)^n$ を級数展開したものを一般式で表すと次のようになります.

$$
(1+x)^n = \sum_{k=0}^{n} \frac{n!}{k!(n-k)!}x^k
$$

(この式の導出については「CG のためのフーリエ解析入門 フーリエ級数編」で説明していますので, そちらを参照してください)

$x = b/a$ を代入すると

$$
\left(1 + \frac{b}{a}\right)^n = \sum_{k=0}^{n} \frac{n!}{k!(n-k)!}\left(\frac{b}{a}\right)^k
$$

これを変形します

$$
\begin{aligned}
\left(\frac{1}{a}\right)^n (a+b)^n &= \sum_{k=0}^{n} \frac{n!}{k!(n-k)!}\left(\frac{b}{a}\right)^k \\
a^{-n}(a+b)^n &= \sum_{k=0}^{n} \frac{n!}{k!(n-k)!}a^{-k}b^k \\
(a+b)^n &= \sum_{k=0}^{n} \frac{n!}{k!(n-k)!}\frac{a^{-k}}{a^{-n}}b^k \\
(a+b)^n &= \sum_{k=0}^{n} \frac{n!}{k!(n-k)!}a^{n-k}b^k
\end{aligned}
$$

組合せの式は

$$
\frac{n!}{r!(n-r)!} = {}_nC_r
$$

なので, 置き換えると

$$
(a+b)^n = \sum_{r=0}^{n} {}_nC_ra^{n-r}b^r
$$

となり, 2 項定理の式となります.

2 項係数を下のように並べたものをパスカルの3角形といいます.

$$
\begin{array}{ccccccccc}
 & & & & {}_0C_0 & & & & \\
 & & & / & & \backslash & & & \\
 & & & {}_1C_0 & & {}_1C_1 & & & \\
 & & / & & \backslash\ / & & \backslash & & \\
 & & {}_2C_0 & & {}_2C_1 & & {}_2C_2 & & \\
 & / & & \backslash\ / & & \backslash\ / & & \backslash & \\
 & {}_3C_0 & & {}_3C_1 & & {}_3C_2 & & {}_3C_3 & \\
/ & & \backslash\ / & & \backslash\ / & & \backslash\ / & & \backslash \\
{}_4C_0 & & {}_4C_1 & & {}_4C_2 & & {}_4C_3 & & {}_4C_4
\end{array}
\quad \text{(a)}
\qquad
\begin{array}{ccccccccc}
 & & & & 1 & & & & \\
 & & & / & & \backslash & & & \\
 & & & 1 & & 1 & & & \\
 & & / & & \backslash\ / & & \backslash & & \\
 & & 1 & & 2 & & 1 & & \\
 & / & & \backslash\ / & & \backslash\ / & & \backslash & \\
 & 1 & & 3 & & 3 & & 1 & \\
/ & & \backslash\ / & & \backslash\ / & & \backslash\ / & & \backslash \\
1 & & 4 & & 6 & & 4 & & 1
\end{array}
\quad \text{(b)}
$$

パスカルの3角形を見ると，任意の2項係数はその上段の2数の和で求めることができます.例えば，${}_3C_1$ は ${}_2C_0 + {}_2C_1 = 1 + 2 = 3$ となっています.これから次のような関係が成り立っていることがわかります.

$$
{}_nC_r = {}_{n-1}C_{r-1} + {}_{n-1}C_r
$$

例題 $\left(x^2 - \frac{2}{x}\right)^6$ の展開

$$
\left(x^2 - \frac{2}{x}\right)^6 = \left\{x^2 + \left(-\frac{2}{x}\right)\right\}^6
$$

として，2項定理をつかって展開します.

$$
\begin{aligned}
\left(x^2 - \frac{2}{x}\right)^6 &= {}_6C_0(x^2)^6 + {}_6C_1(x^2)^5\left(-\frac{2}{x}\right)^1 + {}_6C_2(x^2)^4\left(-\frac{2}{x}\right)^2 + {}_6C_3(x^2)^3\left(-\frac{2}{x}\right)^3 \\
&\quad + {}_6C_4(x^2)^2\left(-\frac{2}{x}\right)^4 + {}_6C_5(x^2)^1\left(-\frac{2}{x}\right)^5 + {}_6C_6\left(-\frac{2}{x}\right)^6 \\
&= x^{12} + 6x^{10}\frac{-2}{x} + 15x^8\frac{4}{x^2} + 20x^6\frac{-8}{x^3} + 15x^4\frac{16}{x^4} + 6x^2\frac{-32}{x^5} + \frac{64}{x^6} \\
&= x^{12} - 12x^9 + 60x^6 - 160x^3 + 240 - \frac{192}{x^3} + \frac{64}{x^6}
\end{aligned}
$$

例題 ${}_nC_0 + {}_nC_1 + \cdots + {}_nC_n$ を求めてください

$$
(1+x)^n = {}_nC_0 + {}_nC_1x + \cdots + {}_nC_rx^r + \cdots + {}_nC_nx^n
$$

に $x = 1$ を代入すると

$$
{}_nC_0 + {}_nC_1 + \cdots + {}_nC_r + \cdots + {}_nC_n = (1+1)^n = 2^n
$$

例題 ${}_nC_0 - {}_nC_1 + {}_nC_2 - \cdots + (-1)^n{}_nC_n$ を求めてください

$$
(1+x)^n = {}_nC_0 + {}_nC_1x + \cdots + {}_nC_rx^r + \cdots + {}_nC_nx^n
$$

に $x = -1$ を代入すると

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

例題 $1 \cdot {}_nC_1 + 2 \cdot {}_nC_1 + 3 \cdot {}_nC_2 + \cdots + n \cdot {}_nC_n$ を求めてください

$$(1+x)^n = {}_nC_0 + {}_nC_1 x + \cdots + {}_nC_r x^r + \cdots + {}_nC_n x^n$$

の両辺を x で微分すると

$$n(1+x)^{n-1} = {}_nC_1 + 2\cdot{}_nC_2 x + 3\cdot{}_nC_3 x^2 + \cdots + n\cdot{}_nC_n x^n$$

これに $x=1$ を代入すると

$${}_nC_1 + 2\cdot{}_nC_2 + 3\cdot{}_nC_3 + \cdots + n\cdot{}_nC_n = n(1+1)^{n-1} = n\cdot 2^{n-1}$$

7. 多項定理

2 項定理では項の数が 2 つでしたが, 発展させて 3 つ以上の場合を考えます.例えば, 3 項式の 3 乗を展開すると次のようになります.

$$(a+b+c)^3 = a^3 + b^3 + c^3 + 3a^2b + 3ab^2 + 3a^2c + 3ac^2 + 3b^c + 3bc^2 + 6abc$$

各項は $a^\alpha b^\beta c^\gamma$ の形をしています.その係数は 3 個の因数 $(a+b+c)$ から a を α 個, b を β 個, c を γ 個選ぶ組合せの数になります.ここで, $\alpha+\beta+\gamma=n$ としたとき, 各項の係数は

$$\frac{n!}{\alpha!\beta!\gamma!}$$

となります.これを一般化して n 乗した場合の展開式は

$$(a+b+c)^n = \sum \frac{n!}{\alpha!\beta!\gamma!} a^\alpha b^\beta c^\gamma \qquad (\alpha+\beta+\gamma=n)$$

となります.さらに m 項式まで拡張すると

$$(a_1+a_2+\cdots+a_n)^n = \sum \frac{n!}{n_1!n_2!\ldots n_m!} a_1^{n_1} a_2^{n_2} \ldots a_m^{n_m}$$

$$(n_1+n_2+\cdots+n_m=n)$$

これを多項定理といいます.ここで $m=2$ としたとき, $n_2=n-n_1$ となって, 係数は $\frac{n!}{n_1!(n-n_1)!} = {}_nC_{n_1}$ となって 2 項定理と一致します.

例題 $(a+b+c)^8$ の展開式には, 何種類の項がありますか

展開式の項は $a^\alpha b^\beta c^\gamma (\alpha+\beta+\gamma=n)$ なので, a,b,c の 3 文字のうちから重複を許して 8 文字とった積となります.よって

$${}_3H_8 = {}_{3+8-1}C_8 = {}_{10}C_8 = {}_{10}C_2 = \frac{10\times 9}{2\times 1} = 45$$

となります.

例題 $(x+y+2z)^6$ の展開式で, x^3y^2z の係数はいくつになりますか

多項定理より x^3y^2z の一般項は次のようになります.

$$\frac{6!}{n_1!n_2!n_3!}x^{n_1}y^{n_2}(2z)^{n_3}$$

$n_1 = 3, n_2 = 2, n_3 = 1$ として x^3y^2z の係数は

$$\frac{6!}{3!2!1!}1^3 1^2 2^1 = \frac{6 \times 5 \times 4}{2} \times 2 = 120$$

となります.

8. 確率

細工をしていないサイコロを何回も振ったときの出る目はそれぞれ偶然によって決まります.このように同じ条件で何度でも繰り返す操作を試行といいます.そして, 試行によって起こる結果を事象といいます.たとえば, サイコロを振る試行の事象はサイコロの目である「1, 2, 3, 4, 5, 6」のことで, 1つひとつを標本といいます.そして, すべての標本(ここでは 1, 2, 3, 4, 5, 6 の6つ)を標本空間といいます.事象は1つの標本でもいいし, 標本空間でも構いません.また, 標本が1つだけの事象を根元事象といいます.

8.1 数学的確率

標本空間 U が n 個の根元事象で構成されていて, どの2つの根元事象も「重複して起こらず」, また, どの根元事象の起こることも「同様に確からしい」とします.ある事象 A が a 個の根元事象から成り立つとき, つまり, A の起こる場合が a 通りであるとき,

$$p = \frac{a}{n},$$

のことを, 事象 A の起こる確率といい, $P(A)$ と表します.これを数学的確率と呼びます.

確率 $P(A) = 0$ のとき, 事象 A は絶対に起こりません. $P(A) = 1$ のときは, 必ず起こります.サイコロの例で1の目が出る確率は, 事象 $A = \{1\}$ で, 標本空間 $U = \{1, 2, 3, 4, 5, 6\}$ なので, 確率は次のようになります.

$$P(A) = \frac{1}{6}.$$

8.2 統計的確率

数学的確率は同様に確からしいという状況のもとに考える確率です.それに対して実際に試行した結果をもとに確率を考える方法があります.たとえば, 野球の打率やゲームで対戦したときの勝率などは何度も試行した結果から計算し, その確からしさを予想します.このような場合, 次のように確率を定義します. n 回試行を行った結果, ある事象 A が r 回起こったとします. n を大きくしていくとき, r/n が一定の値 p に近づくなら, 事象 A が起こる確率 $P(A)$ を

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

とします.この確率を統計的確率といいます.また, r/n を相対頻度といいます.現実的に試行を無限回行うのは不可能ですが, 大数の法則(Law of Large Numbers, またはベルヌーイの定理)によって, n が十分大きいとき, 相対頻度を確率 p としても誤差が小さいことが証明されています.

8.3 確率の性質

集合には和集合や共通集合，補集合というのがありました．確率の事象も集合なので同じことがいえます．事象 A または事象 B が起こる事象を $A \cup B$ と表し，これを和事象といいます．また，事象 A と事象 B がともに起こる事象を $A \cap B$ と表し，これを積事象といいます．事象 A が起こらないという事象を $\overline{A}$ で表し，これを余事象といいます．決して起こらない事象のことを空事象といい，ϕ で表します．

標本空間を U，ある事象 A とすると

$$0 \leq P(A) \leq 1$$

が成り立ちます．これは確率は 0 から 1 の値しかとらないということです．

標本空間と空事象のそれぞれの確率は

$$P(U) = 1, \qquad P(\phi) = 0$$

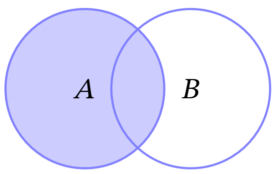
となります．

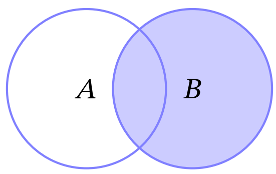
ここで標本空間 $U = \{1, 2, 3, 4, 5, 6\}$ を使って，和事象や積事象との関係を見ていきます．また，ベン図と呼ばれる図を使うと見てわかりやすくなります．

標本空間のうち偶数の事象を A，3 の倍数を B とします．

$$\begin{aligned} A &= \{x = 2m | 2, 4, 6\} = \{2, 4, 6\} \\ B &= \{x = 3m | 3, 6\} = \{3, 6\} \end{aligned}$$

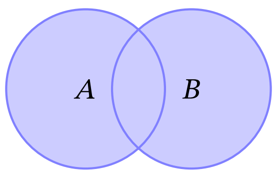
A と B のベン図は次のようになります．





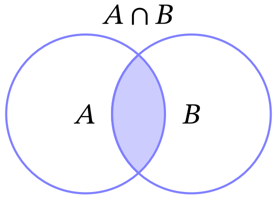
A と B の和事象 $A \cup B$ は $\{2, 3, 4, 6\}$ となります．

$A \cup B$



A と B の積事象 $A \cap B$ は $\{6\}$ となります．

$A \cap B$



積事象 $A \cap B$ の確率はベン図から

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

となります.また,

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

も成立します.もし A と B が互いに排反なら積事象は空事象となるので

$$P(A \cup B) = P(A) + P(B)$$

となります.

次に余事象の起こる確率について, ある事象 E の余事象 $\overline{E}$ を考えると

$$E \cup \overline{E} = U \qquad E \cap \overline{E} = \phi$$

なので

$$1 = P(U) = P(E \cup \overline{E}) = P(E) + P(\overline{E})$$

よって

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

となります.

$\boxed{\text{例題}}$ サイコロを2個同時に振るとき, 出た目の和が5になる確率 p はいくつになりますか

サイコロは6通りの目があるので, 起こり得る場合の数は $6 \times 6 = 36$ です.出た目の和が 5 となるのは, 2 つのサイコロの目が $(1,4), (2,3), (3,2), (4,1)$ の 4 通りなので, 求める確率は

$$p = \frac{4}{36} = \frac{1}{9}$$

となります.

$\boxed{\text{例題}}$ 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 の番号が付いている玉がそれぞれ1つずつ, 合計 10 個あります.この中から 2 個の玉を取り出すとき, (a)2 個の玉に付いている番号の和が 12 になる確率と (b)2 個の玉に付いている番号の和が 6 の倍数である確率はいくつになりますか

まず, 起こりうる場合の総数は

$${}_{10}C_2 = 45$$

です. (a) の和が 12 となるのは $(2,10), (9,3), (8,4), (7,5), (6,6)$ の 5 通りなので

$$p = \frac{5}{45} = \frac{1}{9}$$

となります. (b) の場合, 和が 6 になる場合は $(1,5),(2,4),(3,3)$ の 3 通り, 和が 12 になる場合は (a) から 5 通り, 和が 18 になる場合は $(9,9),(10,8)$ の 2 通りで, 合わせると $3+5+2=10$ 通り.よって求める確率は

$$p = \frac{10}{45} = \frac{2}{9}$$

となります.

8.4 条件付き確率

事象 A,B の余事象を $\overline{A},\overline{B}$ とします.互いに排反する 4 つの事象の起こる場合の数は右表のようになります.このとき, 起こり得る場合の総数は

	B	$\overline{B}$
A	$A\cap B$ a 通り	$A\cap\overline{B}$ b 通り
$\overline{A}$	$\overline{A}\cap B$ c 通り	$\overline{A}\cap\overline{B}$ d 通り

$$n = a+b+c+d$$

となります.事象 A と事象 B がともに起こる確率 p は

$$p = P(A\cap B) = \frac{a}{n}$$

になります.ここで, 事象 A が起こる確率を p_1, 事象 A が起こったときに B が起こる確率を p_2 とすると

$$p_1 = \frac{a+b}{n}, \qquad p_2 = \frac{a}{a+b}$$

となりますから,

$$p_1 \times p2 = \frac{a+b}{n} \times \frac{a}{a+b} = \frac{a}{n} = p$$

が成り立つことがわかります.この p_2 を A が起こったときの B の条件つき確率といい, $P(B|A)$ または $P_A(B)$ で表します.この記号を使うと

$$P(A\cap B) = P(A)\times P(B|A)$$

と表せます.式を変形して $P(B|A)$ を求めると

$$P(B|A) = \frac{P(A\cap B)}{P(A)}$$

となります.また,

$$P(B|A) = \frac{P(A\cap B)}{P(A)} = \frac{\frac{a}{n}}{\frac{a+b}{n}} = \frac{a}{a+b} = \frac{n(A\cap B)}{n(A)}$$

という関係になります.

事象 A が起こったときに事象 B が起こる確率が求められたので, 逆に事象 B が起こる確率を p_1, 事象 B が起こったときに A が起こる確率を p_2 に置き換えてみると

$$p_1 = \frac{a+c}{n}, \quad p_2 = \frac{a}{a+c}, \quad p_1 \times p2 = \frac{a+c}{n} \times \frac{a}{a+c} = \frac{a}{n} = p$$
$$P(A \cap B) = P(B) \times P(A|B)$$
$$\therefore P(A|B) = \frac{P(A \cap B)}{P(B)}$$

まとめると

$$P(A \cap B) = P(A) \times P(B|A) = P(B) \times P(A|B)$$

となります.これを確率の乗法定理といいます.

例題 10 本中に 3 本の当たりくじがあります.太郎が先に,次郎が後に引いたときにそれぞれ当たりくじを引く確率はいくつになりますか.くじは引いたあと戻さないことにします

太郎が当たりくじをひく事象を A, 次郎が当たりくじを引く事象を B とします.太郎が当たりくじを引く確率は

$$P(A) = \frac{3}{10}$$

また,その余事象 $(\overline{A})$ は

$$P(\overline{A}) = 1 - P(A) = 1 - \frac{3}{10} = \frac{7}{10}$$

次郎が当たりくじを引くのは 2 つの場合があり,太郎が当たりくじを引いたときと引かなかったときです.太郎が当たりくじを引いたとき,次郎が当たりくじを引く確率は

$$P(A \cap B) = P(A)P(B|A) = \frac{3}{10} \times \frac{2}{9} = \frac{6}{90}$$

太郎が当たりくじを引かなったとき,次郎が当たりくじを引く確率は

$$P(\overline{A} \cap B) = P(\overline{A})P(B|A) = \frac{7}{10} \times \frac{3}{9} = \frac{21}{90}$$

事象 $A \cap B$ と $\overline{A} \cap B$ は互いに排反であり, $B = (A \cap B) \cup (\overline{A} \cap B)$ なので

$$P(B) = P(A \cap B) + P(\overline{A} \cap B) = \frac{6}{90} + \frac{21}{90} = \frac{27}{90} = \frac{3}{10}$$

つまり,次郎が後にくじを引いても当たりくじを引く確率は太郎と変わらないことがわかります.

9. 独立試行

条件付き確率 $P(B|A)$ は A が起こったときに, B が起こる確率でした.ですが, A が B の確率に影響しない場合があります.例えば,サイコロを 2 個順番に振った場合, 1 個目に 1 の出目が出る事象を A, 2 個目に 2 の出目が出る事象を B としたとき,起こり得る場合の数は $6 \times 6 = 36$ 通りで, A は $(1,1),(1,2),(1,3),(1,4),(1,5),(1,6)$ の 6 通り, B も $(1,2),(2,2),(3,2),(4,2),(5,2),(6,2)$ の 6 通り, A と B がともに起こる場合 $(A \cap B)$ は $(1,2)$ の 1 通りなので

$$P(A)=\frac{6}{36}=\frac{1}{6},\quad P(B)=\frac{6}{36}=\frac{1}{6},\quad P(A\cap B)=\frac{1}{36}$$
$$P(B|A)=\frac{P(A\cap B)}{P(A)}=\frac{\frac{1}{36}}{\frac{1}{6}}=\frac{1}{6}$$
$$\therefore P(B|A)=P(B)$$

条件つき確率 $P(B|A)$ と B が起こる確率 $P(B)$ が同じなので，B が起こる確率に A が影響していないことがわかります.このような関係から事象 A と B は独立であるといいます.ここで，10 本の当たりくじのときを考えてみましょう.太郎と次郎がともに当たりくじを引いたときの条件つき確率 $P(B|A)$ と次郎が当たりくじを引く確率 $P(B)$ は次のとおりです.

$$P(B|A)=\frac{2}{9},\quad P(B)=\frac{3}{10}$$
$$\therefore P(B|A)\neq P(B)$$

この場合は太郎と次郎がともに当たりくじを引く確率と次郎が当たりくじを引く確率が一致しません.このような関係から事象 A と B は従属であるといいます.

これらから，事象 A と B が独立のとき $P(B|A)=P(B)$，従属のとき $P(B|A)\neq P(B)$ が成り立ちます.また，式を変形して

$$P(B|A)=P(B)$$
$$\frac{P(A\cap B)}{P(A)}=P(B)$$
$$P(A\cap B)=P(A)\times P(B)$$

という関係になります.

事象 A と B が独立であるならば，A によって B の確率は変わりません.また，その逆も然りです.例えば，サイコロのように 1 回ごとの試行が他の試行に影響しないならば，この試行は独立していると言えそうです.まさにこのような試行を独立試行といいます.また，ベルヌーイ試行ともいいます.

この独立試行を何回か繰り返したとき，ある事象が何回起こるかについての確率を考えます.

n 回の独立試行のうち，事象 A の起こる回数が r で，それが何回目に起きるかは ${}_nC_r$ 通りあります.その 1 つの場合の確率を考えると，事象 A が r 回起こり，事象 A の余事象が $n-r$ 回起こるので，事象 A の起こる確率を p，余事象の起こる確率を $q=1-p$ とすれば，乗法定理より

$$p^r(1-p)^{n-r}=p^rq^{n-r}$$

このような確率をもつ場合が ${}_nC_r$ 通りあり，それらは互いに排反であるから，事象 A がちょうど r 回起こる確率 p_r は

$$p_r={}_nC_rp^rq^{n-r}$$

となります.ここで，2 項定理を思い出してみましょう.

$$(a+b)^n=\sum_{r=0}^{n}{}_nC_ra^{n-r}b^r$$

さきほどの式は 2 項定理の一般項と一致します.独立試行と 2 項定理は密接な関係がありそうです.それについては 2 項分布のところで触れていきます.

例題 サイコロを 6 回振って, そのうち 4 回, 1 の目が出る確率はいくつになりますか

各試行において, 1 の目が出る確率は $\frac{1}{6}$ なので, 求める確率は

$$p = {}_6C_4 \left(\frac{1}{6}\right)^4 \left(\frac{5}{6}\right)^2 = \frac{6 \times 5}{2} \frac{5^2}{6^6} = \frac{15 \times 25}{6^6} = \frac{375}{46656} = \frac{125}{15552}$$

10. ベイズの定理

例えば, プログラマが 3 人いました.彼らは単純なモジュールを量産していて, 彼らが作ったモジュールであるという事象をそれぞれ A,B,C で表します.また, 作成されたモジュールはそれぞれ $P(A), P(B), P(C)$ の割合で作られています.彼らはときどきバグがあるモジュールを作成することがあり, そのようなモジュールは不良モジュールとなります.作成されたモジュールが不良モジュールであるという事象を E としたとき, 彼らが不良モジュールを作る確率を $P(E|A), P(E|B), P(E|C)$ とし, この確率がわかっているとします.このとき, 実際に不良モジュールが見つかったときに, 誰がその不良モジュールを作ったかどうかの確率 $P(A|E), P(B|E), P(C|E)$ を考えます.まず, $P(A)$ など, 不良モジュールが作成されたときと関係がないのでこのような確率を事前確率といいます.また, $P(A|E)$ など不良モジュールが作成されたときの条件つき確率なので, これを事後確率といいます.それでは, 不良モジュールが作成されたときに, それが A であるかの確率 $P(A|E)$ を求めます.乗法定理から

$$P(A)P(E|A) = P(E)P(A|E)$$

であるから, 式を変形すると

$$P(A|E) = \frac{P(A)P(E|A)}{P(E)}$$

事象 A, B, C はそれぞれ排反なので $E = (A \cap E) \cup (B \cap E) \cup (C \cap E)$ だから,

$$P(E) = P(A \cap E) + P(B \cap E) + P(C \cap E)$$

乗法定理から

$$P(A \cap E) = P(A)P(E|A)$$

が成り立つので

$$P(E) = P(A)P(E|A) + P(B)P(E|B) + P(C)P(E|C)$$

となります.よって,

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

となります.これを一般の n 事象に拡張することができます.ある事象 E が, n 個の互いに排反である事象 $A_1, A_2, \dots, A_n$ によっているとき, かつ次の条件を満たしているとき

$$E = \sum_{k=1}^{n} E \cap A_k$$

ある事象 A_i によって起こる確率 $P(A_i|E)$ は

$$P(A_i|E) = \frac{P(A_i)P(E|A_i)}{\sum_{k=1}^{n} P(A_k)P(E|A_k)}$$

となります.これをベイズの定理といいます.

例題 ある生産工場で 3 つの機械 (A, B, C) の生産量と, 不良品の出る割合は表の通りです. 1 つの生産品を取り出したら, それが不良品であったとき, A が生産したものである確率はいくつになりますか

機械	生産量の割合	不良品の割合
A	10%	4%
B	40%	2%
C	50%	1%

それぞれの確率を求めると

$$P(A) = 0.1, \quad P(B) = 0.4, \quad P(C) = 0.5$$
$$P(E|A) = 0.04, \quad P(E|B) = 0.02, \quad P(E|C) = 0.01$$

よって

$$\begin{aligned} P(A|E) &= \frac{0.1 \times 0.04}{0.1 \times 0.04 + 0.4 \times 0.02 + 0.5 \times 0.01} \\ &= \frac{0.004}{0.004 + 0.008 + 0.005} = \frac{0.004}{0.017} \cong 0.235 \\ &\cong 24\% \end{aligned}$$

11. ポーカーゲームの確率

確率計算の練習としてポーカーゲームの確率を求めてみます.ポーカーの役は, ワンペア, ツーペア, スリーカード, ストレート, フラッシュ, フルハウス, フォーカード, ストレートフラッシュ, ロイヤルストレートフラッシュがあります.ここでは, ジョーカーを含めず, 最初に5枚配られたときのそれぞれの確率を求めます.

まず, ポーカーでは, ジョーカーを除いた 52 枚のカードから 5 枚のカードを選びます.これが起こり得るすべての場合の数となります.

$${}_{52}C_5 = \frac{52!}{47!5!} = \frac{52 \times 51 \times 50 \times 49 \times 48}{5 \times 4 \times 3 \times 2} = 2598960$$

11.1 ワンペア

トランプのカードは 1~から 13 までの数字と4種類のスート(ダイヤ, ハート, スペード, クラブ) があります.ワンペアは同じ数字のカードが 2 枚 1 組なので, まずペアになる数字を選ぶ方法を考えると ${}_{13}C_1$ 通りです.残りの 3 枚はばらばらの数字になっている必要があるので, ${}_{12}C_3$ 通りになります.よって, 数字の選び方は

$${}_{13}C_1 \times {}_{12}C_3 = \frac{13!}{(13-1)!1!} \times \frac{12!}{(12-3)!3!} = 13 \times \frac{12 \times 11 \times 10}{3 \times 2} = 2860$$

次に, スートを選ぶ方法を考えます.ワンペアのカードについては 4 種類から 2 種類選ぶ方法となり, ${}_{4}C_2$ 通りです.残りのカード 3 枚についてはそれぞれ 4 種類から 1 種類を選ぶので ${}_{4}C_1$ となります.よって, スートの選び方は

$${}_{4}C_2 \times {}_{4}C_1 \times {}_{4}C_1 \times {}_{4}C_1 = \frac{4!}{2!2!} \times 4 \times 4 \times 4 = 384$$

ワンペアの場合の数は数字とスートの選び方の積となって

$$2860 \times 384 = 1098240$$

したがって, ワンペアの出る確率は

$$\frac{1098240}{2598960} \cong 0.4225$$

11.2 ツーペア

ツーペアは同じ数字のカードが 2 枚 2 組となります.まず, ツーペアとなる数字の選び方は ${}_{13}C_2$ 通りです.また, 残りの数字は 11 枚の中から 1 枚を選ぶので ${}_{11}C_1$ 通りです.よって, 数字の選び方は

$${}_{13}C_2 \times {}_{11}C_1 = \frac{13!}{(13-2)!2!} \times \frac{11!}{(11-1)!1!} = \frac{13 \times 12}{2 \times 1}\frac{1}{1} = 858$$

次にスートについてですが, ツーペアのカードはそれぞれ 2 種類選ぶので ${}_{4}C_2$ 通りとなります.残りの 1 枚は 4 種類から 1 種類を選ぶので ${}_{4}C_1$ 通りです.よって

$${}_{4}C_2 \times {}_{4}C_2 \times {}_{4}C_1 = \frac{4!}{2!2!} \times \frac{4!}{2!2!} \times 4 = \frac{4 \times 3}{2 \times 1} \times \frac{4 \times 3}{2 \times 1} \times 4 = 144$$

ツーペアの場合の数は

$$858 \times 144 = 123552$$

と与えられ, ツーペアの出る確率は

$$\frac{123552}{2598960} \cong 0.0475$$

11.3 スリーカード

スリーカードは同じ数字のカードが 3 枚と異なる 2 枚のカードになります.スリーカードとなる数字の選び方は ${}_{13}C_1$ 通りです.残りの 2 枚のカードの数字の選び方は ${}_{12}C_2$ 通りです.よって

$${}_{13}C_1 \times {}_{12}C_2 = 13 \times \frac{12!}{(12-2)!2!} = 13 \times \frac{12 \times 11}{2 \times 1} = 858$$

次にスリーカードとなるカードのスートの選び方は ${}_4C_3$ 通りで, 残り 2 枚のカードのスートの選び方はそれぞれ ${}_4C_1$ 通りとなり

$${}_4C_3 \times {}_4C_1 \times {}_4C_1 = \frac{4!}{1!3!} \times 4 \times 4 = 64$$

スリーカードの場合の数は

$$858 \times 64 = 54912$$

と与えられ, スリーカードの出る確率は

$$\frac{54912}{2598960} \cong 0.021$$

11.4 ストレート

ストレートは 5 枚のカードが連続したものです.例えば JQKA2 というようにキングとエースをまたぐことはできません.この場合, 先頭がエースになる場合から 10 までの 10 通りとなります.次にスートは 5 枚それぞれ 4 種類から 1 種類選ぶので ${}_4C_1$ 通りです.ただし, 5 枚すべて同じスートの場合はストレートフラッシュ(ロイヤルストレートフラッシュ含む)となるので, その分を差し引きます.それは数字の選び方が 10 通りでスート 4 種類分となるので 40 通りとなって

$$10 \times 4 \times 4 \times 4 \times 4 \times 4 - 40 = 10200$$

よって, ストレートの出る確率は

$$\frac{10200}{2598960} \cong 0.0039$$

11.5 フラッシュ

フラッシュはすべて同じスートになっているものです.この場合, 5 枚のカードは数字はなんでもいいので ${}_{13}C_5$ 通り, スートは 4 種類から 1 種類なので 4 通り.ただし, これもストレートフラッシュ(ロイヤルストレートフラッシュ含む)を差し引きます.よって

$${}_{13}C_5 \times 4 = \frac{13 \times 12 \times 11 \times 10 \times 9}{5 \times 4 \times 3 \times 2} \times 4 - 40 = 5108$$

よって, フラッシュの出る確率は

$$\frac{5108}{2598960} \cong 0.0019$$

11.6 フルハウス

フルハウスはスリーカードとワンペアの組合せです.スリーカードになる数字の選び方は ${}_{13}C_1$ 通りで, ワンペアになる数字の選び方は ${}_{12}C_1$ 通りとなって

$$ {}_{13}C_1 \times {}_{12}C_1 = 13 \times 12 = 156 $$

つぎに, スリーカードとなるカードのスートは 4 種類から 3 種類選ぶので ${}_4C_3$ 通りで, ワンペアでは 2 種類選ぶので ${}_4C_2$ 通り.よって

$$ {}_4C_3 \times {}_4C_2 = \frac{4!}{1!3!} \times \frac{4!}{2!2!} = 4 \times 6 = 24 $$

フルハウスの場合の数は

$$ 156 \times 24 = 3744 $$

と与えられ, フルハウスの出る確率は

$$ \frac{3744}{2598960} \cong 0.0014 $$

11.7 フォーカード

フォーカードは同じ数字のカードが 4 枚と異なる 1 枚のカードになります.フォーカードとなる数字の選び方は ${}_{13}C_1$ 通りです.残りの 1 枚のカードの数字の選び方は ${}_{12}C_1$ 通りです.よって

$$ {}_{13}C_1 \times {}_{12}C_1 = 13 \times 12 = 156 $$

次にフォーカードとなるカードのスートは 4 枚全部なので 1 通り, 残り 1 枚のカードのスートの選び方は ${}_4C_1$ 通りとなり

$$ 1 \times {}_4C_1 = 4 $$

フォーカードの場合の数は

$$ 156 \times 4 = 624 $$

と与えられ, フォーカードの出る確率は

$$ \frac{624}{2598960} \cong 0.00024 $$

11.8 ストレートフラッシュ

ストレートフラッシュはストレートとフラッシュの組合せです.ストレートのときに差し引いた分で, 40 通りとなります.さらに, ロイヤルストレートフラッシュの組合せを差し引くと, $40 - 4$ となって, 36 通りとなります.

ストレートフラッシュの出る確率は

$$ \frac{36}{2598960} \cong 0.0000138 $$

11.9 ロイヤルストレートフラッシュ

ロイヤルストレートフラッシュはエース, キング, クイーン, ジャック, 10 の組合せでストレートフラッシュのものです.数字の選び方は 1 通りで, スートの選び方は 4 種類なので 4 通りとなります.よって, ロイヤルストレートフラッシュの出る確率は

$$\frac{4}{2598960} \cong 0.0000015$$

11.10 ノーペア

すべての役を足したものの余事象となるので, 場合の数は

$$1098240 + 123552 + 54912 + 10200 + 5108 + 3744 + 624 + 36 + 4 = 1296420$$
$$2598960 - 1296420 = 1302540$$

よって, 何も役がない確率は

$$\frac{1302540}{2598960} \cong 0.5$$

11.11 ポーカーの役の確率まとめ

ポーカーの役の確率を表にすると次のようになります.

ノーペア	約 50%
ワンペア	約 42.25%
ツーペア	約 4.75%
スリーカード	約 2.1%
ストレート	約 0.39%
フラッシュ	約 0.19%
フルハウス	約 0.14%
フォーカード	約 0.024%
ストレートフラッシュ	約 0.0014%
ロイヤルストレートフラッシュ	約 0.00015 %

12. 確率分布

これまで, ある事象が起こる確率を求めてきました.事象ごとの確率を取り扱うときに, 事象に対応した実数を変数として, その変数に対応した確率が与えられるようにします.このような変数を確率変数といい, X, Y などの大文字で表します.ある試行の結果(つまり事象) ω によって X の値が決まる場合, $X(\omega)$ とすれば, その事象が起こる確率が与えられます.例えば, サイコロの出目を確率変数とすると, X は $\{1, 2, 3, 4, 5, 6\}$ の値をとれることになります.ここで, 確率変数とその確率を表にまとめると次のようになります.

確率変数 X	1	2	3	4	5	6
確率 $P(X)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

このような表を確率分布表といいます.確率分布表を見るとすべての事象に対して確率がどのように分布しているかがわかります.このサイコロの例で見ると確率が均等になっています.例えば確率変数が 1 のときの確率は

$$P(X=1)=\frac{1}{6}$$

と表記することができます.

確率変数は実数の値をとりますが,たとえばコインの表裏など直接数値に対応していない場合があります.その場合は表を 1 ,裏を 0 として対応させせます.すると

$$P(X=1)=\frac{1}{2},\quad P(X=0)=\frac{1}{2}$$

となります.確率変数はサイコロの出目やコインの表裏などのように離散の値や,連続した値をとることができます.例えば,確率変数が身長の値をとるとき, $P(170.1\leq X\leq 175.8)$ とすれば 170.1~175.8 の確率が得られます.

離散の確率変数に対して, X が $x_i(i=1,2,\ldots,n)$ となる確率を

$$P(X=x_i)=p_i$$

という関係にあるとき,確率は関数の形で

$$f(x)=\begin{cases} p_i & (x=x_i) \\ 0 & (x\neq x_i) \end{cases}$$

と書けます.このような関数を確率密度関数(**probability density function:pdf**)といいます.すべての確率の和は 1 なので

$$\sum_{i=1}^{n} f(x_i)=1 \qquad f(x_i)\geq 0$$

が成り立ちます.また,確率変数 X のとる値が x 以下である確率について

$$F(x)=P(X\leq x)$$

という関数を考えることができます.このような関数を累積分布関数 (**cumulative distribution function:cdf**)といいます.

例えば,サイコロを振ったときの確率密度関数と累積分布関数のグラフは次のようになります.

(a) 確率密度関数　　**(b)** 確率分布関数

Fig.1: 離散の確率密度と分布関数

次に確率変数が連続の場合,離散と同じように確率密度関数と分布関数を考えることができます.確率変数の値の範囲が $a\leq x\leq b$ だとすると, X が x と $x+\Delta x$ の間にある確率は

$$P(x \leq X \leq x + \varDelta x) = \int_{x}^{x+\varDelta x} f(u)du$$

となるような関数 $f(x)$ が連続の確率密度関数です.ここで $\varDelta x$ が小さいとき, 積分は面積で近似できるので

$$P(x \leq X \leq x + \varDelta x) \fallingdotseq f(x)\varDelta x$$

と書くことができます.連続でもすべての確率の和は 1 であるので,

$$\int_{a}^{b} f(x)dx = 1$$

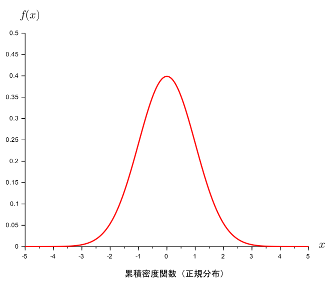
が成り立ちます.また, 区間 $a \leq x \leq b$ に含まれていない場合は $f(x) = 0$ として,

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

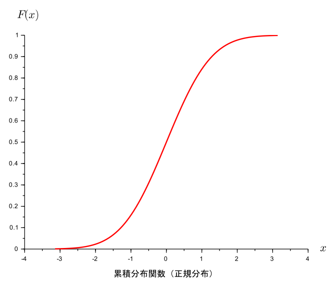
と書くこともできます.次に, 連続の累積分布関数は次のようになります.

$$F(x) = \int_{-\infty}^{x} f(u)du$$

連続の確率密度関数と累積分布関数の例として正規分布(後で詳しく出てきます)のグラフは次のようになります.



(a) 確率密度関数



(b) 累積分布関数

Fig.2: 連続の確率密度と分布関数

離散の確率密度関数と連続の確率分布関数の違いを確認してみると, 離散の確率密度関数では, ある確率変数の確率が

$$p(X = x_i) = f(x_i)$$

の関係になりますが, 連続の場合

$$p(a \leq x \leq a) = p(x = a) = \int_{a}^{a} f(x)dx = 0$$

となって

$$p(x = a) \neq f(a)$$

ということになります.また, 離散と連続の累積分布関数 $F(x)$ を微分すると $f(x)$ が得られます.

$$f(x) = \frac{dF(x)}{dx}$$

これはグラフ上の累積分布関数 $F(x)$ の曲線上の各点での勾配が, 確率密度関数 $f(x)$ の値になっていることを意味しています.

ここで次のような確率密度関数を考えてみます.

$$f(x) = \begin{cases} c & (|x| \leq 1) \\ 0 & (|x| > 1) \end{cases}$$

この分布は確率変数 X が-1 から 1 までの範囲で確率密度が一定(c)であり, 特に一様分布といいます.確率密度関数は

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

を満たす必要があるため

$$\int_{-\infty}^{\infty} f(x)dx = \int_{-1}^{1} cdx = [cx]_{-1}^{1} = c - (-c) = 2c = 1$$

$$\therefore c = \frac{1}{2}$$

区間 $-1 < x < 1$ のとき, $f(x) = 1/2$ なので, 累積分布関数は

$$F(x) = \int_{-\infty}^{x} f(y)dy = \int_{-1}^{x} \frac{1}{2}dy = \left[\frac{1}{2}y\right]_{-1}^{x} = \frac{1}{2}x - \left(-\frac{1}{2}\right) = \frac{1}{2}(x+1)$$

となります. $-1 < x < 1$ の区間で $f(x) = 1/2$ を積分すれば

$$F(x) = \int_{-1}^{1} \frac{1}{2}dx = \left[\frac{1}{2}\right]_{-1}^{1} = \frac{1}{2} - \left(\frac{1}{2}\right) = 1$$

となって確率が 1 となっています.この確率密度関数と累積分布関数のグラフは次のようになります.

(a) 確率密度関数　　**(b)** 累積分布関数

Fig.3: 一様分布の確率密度と分布関数

この一様分布では区間が $-1 < x < 1$ なので, 任意の区間 $a \leq x \leq b$ を考えてみます.つまり,

$$f(x) = c \qquad (a \leq x \leq b)$$

まず, 確率密度関数は負の値にならないので

$$f(x) = c \geq 0$$

となって

$$\int_{-\infty}^{\infty} f(x)dx = \int_{a}^{b} cdx = [cx]_a^b = c(b-a) = 1$$

より

$$c = \frac{1}{b-a} \qquad f(x) = \frac{1}{b-a}$$

となります.今度は次のような確率密度関数を考えてみます.

$$f(x) = \begin{cases} ce^{-x} & (x \geq 0) \\ 0 & (x < 0) \end{cases}$$

まず,確率密度関数は負の値にならないので

$$f(x) = ce^{-x} \geq 0 \qquad \therefore c \geq 0$$

となります.次に

$$\int_{-\infty}^{\infty} fxdx = \int_{0}^{\infty} ce^{-x}dx = [-ce^{-x}]_0^{\infty}$$
$$= \lim_{x\to\infty}(-ce^{-x}) - \lim_{x\to 0}(-ce^{-x}) = 0-(-c) = c = 1$$

よって

$$f(x) = e^{-x}$$

となります.これをより一般的にした場合

$$f(x) = \begin{cases} ce^{-\lambda x} & (x \geq 0) \\ 0 & (x < 0) \end{cases}$$

となって, c を求めると

$$\int_{0}^{\infty} ce^{-\lambda x}dx = \left[\frac{ce^{-\lambda x}}{-\lambda}\right]_0^{\infty} = \frac{c}{\lambda} = 1$$
$$\therefore c = \lambda$$

したがって,この確率密度関数は

$$f(x) = \lambda e^{-\lambda x}$$

そして,累積分布関数は

$$F(x) = \int_{-\infty}^{x} f(y)dy = \int_{0}^{x} \lambda e^{-\lambda y}dy = \left[\frac{\lambda e^{-\lambda y}}{-\lambda}\right]_0^x = \left[-e^{-\lambda y}\right]_0^x = 1 - e^{-\lambda x}$$

となります.このような確率分布を指数分布といいます.このグラフは次のようになります.

(a) 確率密度関数　　**(b)** 累積分布関数

Fig.4: 指数分布の確率密度と分布関数

13. 期待値と分散

確率密度関数がわかれば確率変数に対応した確率が求まりました.また,確率変数とその確率がわかれば,確率の分布がわかるようになります.確率分布を代表する値としてその平均値 μ を考えると,離散の場合

$$\mu = \sum_{i=1}^{n} x_i f(x_i)$$

連続の場合は

$$\mu = \int_{-\infty}^{\infty} x f(x) dx$$

となります.この平均値を特に期待値といい,確率変数 X の期待値を $E[X]$ と表します.つまり $\mu = E[X]$ です.以降は μ も $E[X]$ と同じ期待値として扱います.

例えば,サイコロの出目を確率変数とした場合, $x_1 = 1, x_2 = 2, \ldots, x_6 = 6$ で, $f(x_i) = 1/6$ なので,その期待値は

$$\mu = 1 \times \frac{1}{6} + 2 \times \frac{1}{6} + 3 \times \frac{1}{6} + 4 \times \frac{1}{6} + 5 \times \frac{1}{6} + 6 \times \frac{1}{6} = \frac{21}{6} = 3.5$$

となります.また,サイコロの出目の平均値 $\overline{x}$ を求めると

$$\overline{x} = \frac{1+2+3+4+5+6}{6} = 3.5$$

となって,期待値と一致します.つまり, x にそれが出る確率をかけて足したものは, x の平均値となります.

$$\overline{x} = E[X]$$

次に宝くじの場合を考えてみます.この宝くじは 1 枚 100 円で, 1 等が 1 万円, 2 等が 1000 円, 3 等が 200 円とします.全部で 1000 枚の宝くじがあり, 1 等は 1 枚, 2 等が 10 枚, 3 等は 100 枚あるとします.1 等, 2 等, 3 等の当たる確率はそれぞれ

$$\frac{1}{1000},\quad \frac{10}{1000}=\frac{1}{100},\quad \frac{100}{1000}=\frac{1}{10}$$

となります.この期待値を求めると

$$\mu = 10000\times\frac{1}{1000}+1000\times\frac{1}{100}+200\times\frac{1}{10}=10+10+20=40$$

となって,この宝くじは一本あたり 40 円期待できると考えられます.何枚も宝くじを購入してもこの期待値は変わりませんが,宝くじが当たる確率は上がっていきます.

それでは次の確率密度関数の期待値を求めてみます.

$$f(x)=\begin{cases} xe^{x} & (x\geq 0)\\ 0 & (x<0)\end{cases}$$

この確率変数は連続で指数分布に従っています.よって期待値は

$$\mu=\int_{-\infty}^{\infty} xf(x)dx=\int_{0}^{\infty} x^2e^{-x}dx$$

となります.これは部分積分

$$\int_a^b f(x)g'(x)dx=[f(x)g(x)]_a^b-\int_a^b f'(x)g(x)dx$$

を利用して,$f(x)=x^2, g'(x)=e^{-x}$ とすると $f'(x)=2x, g(x)=-e^{-x}$ だから

$$\begin{aligned}\mu&=\int_0^\infty x^2e^{-x}dx\\ &=\left[x^2(-e^{-x})\right]_0^\infty-\int_0^\infty 2x(-e^{-x})dx\\ &=(0-0)-2\int_0^\infty x(-e^{-x})dx\\ &=2\int_0^\infty x(e^{-x})dx\end{aligned}$$

もう一度部分積分を使って

$$\begin{aligned}\mu&=2\left([x(-e^{-x})]_0^\infty-\int_0^\infty -e^{-x}dx\right)\\ &=2\int_0^\infty e^{-x}dx\\ &=2\left[-e^{-x}\right]_0^\infty=2\times\{0-(-1)\}\\ &=2\end{aligned}$$

となります.この確率密度関数は図 5 のようなグラフになります.

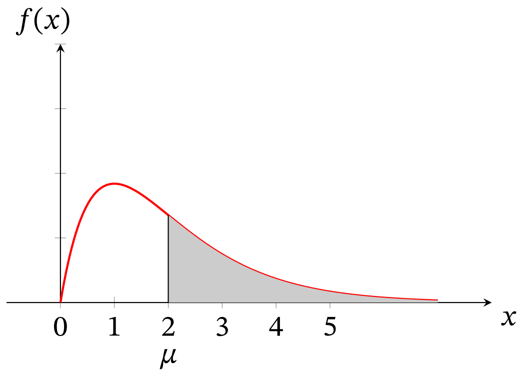


Fig.5: 指数分布の期待値

ここで次のような関数を考えてみます.

$$f(x) = \begin{cases} \displaystyle\int_0^\infty t^{x-1}e^{-t}dt & (x \geq 0) \\ 0 & (x < 0) \end{cases}$$

x に $n+1$ を代入して, 部分積分を利用すると

$$\begin{aligned} f(x) &= \int_0^\infty t^{x-1}e^{-t}dt \\ f(n+1) &= \int_0^\infty t^{n}e^{-t}dt \\ &= \left[-t^n e^{-t}\right]_0^\infty + n\int_0^\infty t^{n-1}e^{-t}dt \\ &= n\int_0^\infty t^{n-1}e^{-t}dt \\ &= nf(n) \end{aligned}$$

となって

$$f(n+1) = nf(n)$$

という漸化式が得られます.この関数をガンマ関数といい, $\Gamma(x)$ と書きます.ガンマ関数は

$$\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}dt$$

でしたので, $x = 1$ とすると

$$\Gamma(1) = \int_0^\infty e^{-t}dt = \left[-e^{t}\right]_0^\infty = 1$$

となります.ガンマ関数は

$$\Gamma(n+1) = n\Gamma(n)$$

の関係から x に正の整数を選ぶと

$$
\begin{aligned}
\Gamma(2) &= 1\Gamma(1) = 1 \\
\Gamma(3) &= 2\Gamma(2) = 2 \cdot 1 \\
\Gamma(4) &= 3\Gamma(3) = 3 \cdot 2 \cdot 1
\end{aligned}
$$

となって

$$\Gamma(n+1) = n \times (n-1) \times (n-2) \times \cdots \times 3 \times 2 \times 1 = n!$$

と階乗になります.そのため,ガンマ関数のことを階乗関数とも呼びます.

話を少し戻すと,確率密度関数が xe^x である指数分布の期待値は

$$\mu = \int_0^\infty x^2 e^{-x} dx = 2$$

でした.これは

$$\Gamma(3) = \int_0^\infty t^2 e^{-t} dt$$

と一致するので

$$\mu = \Gamma(3) = 2! = 2$$

となります.計算が簡単になりますね.せっかくなので,ガンマ関数についてもう少し見ていきましょう.例えばガンマ関数で実数 $x = 1/2$ とすると

$$\Gamma\left(\frac{1}{2}\right) = \int_0^\infty t^{\frac{1}{2}} e^{-t} dt$$

ここで $t = u^2$ とおくと $dt = 2udu$ なので

$$\Gamma\left(\frac{1}{2}\right) = 2\int_0^\infty \frac{u}{\sqrt{u^2}} e^{-u^2} du = 2\int_0^\infty e^{-u^2} du$$

と変形できます.この積分はガウス積分(後で詳しく出てきます)なので

$$\int_0^\infty e^{-u^2} du = \frac{\sqrt{\pi}}{2}$$

となります.よって

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$$

となります.ガンマ関数の漸化式

$$\Gamma(n+1) = n\Gamma(n)$$

を利用すれば

$$\Gamma\left(\frac{3}{2}\right) = \Gamma\left(\frac{1}{2}+1\right) = \frac{1}{2}\Gamma\left(\frac{1}{2}\right) = \frac{\sqrt{\pi}}{2}$$

と計算することができます.また, n を正の整数とすれば

$$\Gamma\left(n+\frac{1}{2}\right) = \frac{(2n)!}{2^{2n}n!}\sqrt{\pi}$$

の関係が成り立ちます.

ガンマ関数として関係して, ベータ関数というのがあります.ベータ関数は $B(\alpha,\beta)$ で表し

$$B(\alpha,\beta) = \int_0^1 x^{a-1}(1-x)^{\beta-1}dx \quad (\alpha>0,\beta>0)$$

と定義されます.ベータ関数とガンマ関数は

$$B(\alpha,\beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$$

という関係になっています.ここで, α,β ともに正の整数 n,m であるとき

$$\begin{aligned}\frac{1}{B(\alpha,\beta)} &= \frac{\Gamma(n+m)}{\Gamma(n)\Gamma(m)} = \frac{(n+m-1)!}{(n-1)!(m-1)!}\\ &= n_{n+m-1}C_{m-1} = m_{n+m-1}C_{n-1}\end{aligned}$$

と書くことができます.ガンマ関数では階乗, ベータ関数の逆数は組合せを整数ではないものに拡張したものと考えることもできます.

ガンマ関数とベータ関数はこれぐらいにして次に進みましょう.

確率変数とその確率分布から期待値が求まりました.その期待値と確率変数 X との差を偏差といいます.式で表すと $X-\mu$ となります.次に偏差の総和を求めてみます.例えば, サイコロの出目で考えると期待値は 3.5 だったので

$$\begin{gathered}(1-3.5)+(2-3.5)+(3-3.5)+(4-3.5)+(5-3.5)+(6-3.5)=\\(-2.5)+(-1.5)+(-0.5)+0.5+1.5+2.5=0\end{gathered}$$

と 0 になってしまいます.そこで偏差の平方 $(X-\mu)^2$ の総和を求めると

$$\begin{gathered}(1-3.5)^2+(2-3.5)^2+(3-3.5)^2+(4-3.5)^2+(5-3.5)^2+(6-3.5)^2=\\6.25+2.25+0.25+0.25+2.25+6.25=17.5\end{gathered}$$

となって意味のある値となりました.この偏差の平方の期待値を求めてみると

$$E[(X-\mu)^2] = \sum_{i=1}^{n}(x_i-\mu)^2 f(x_i)$$

となります.これは離散のときで, 連続の場合は

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

となります.この値を分散といい, 確率変数 X の分散を $V[X]$ と書きます.よって

$$V[X] = E[(X-\mu)^2]$$

となります.また, 分散の平方根を標準偏差といい, σ と書きます.すると

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

という関係になって, σ^2 は分散ということになります.分散は偏差の平方の期待値つまり平均なので, 確率変数 X が平均を中心に, どれくらいばらついているかを表しています.図 6 は分散のばらつきの関係を表しています.

Fig.6: 分散のばらつき

分散は

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

なので, 右辺を展開すると

$$\begin{aligned} E[(X-\mu)^2] &= \int_{-\infty}^{\infty} (x^2 - 2x\mu + \mu^2) f(x)dx \\ &= \int_{-\infty}^{\infty} x^2 f(x)dx - 2\mu \int_{-\infty}^{\infty} x f(x)dx + \mu^2 \int_{-\infty}^{\infty} f(x)dx \\ &= E[x^2] - 2\mu E[x] + \mu^2 \\ &= E[x^2] - 2\mu^2 + \mu^2 \\ &= E[x^2] - \mu^2 \end{aligned}$$

よって

$$E[(X-\mu)^2] = E[X^2] - \mu^2 = E[X^2] - (E[X])^2$$

という関係が成り立ちます.

それではいくつかの分散を求めてみます.まずはサイコロ振りの場合, $\mu = 3.5$ なので

$$
\begin{aligned}
\sigma^2 &= (1-3.5)^2 \times \frac{1}{6} + (2-3.5)^2 \times \frac{2}{6} + (3-3.5)^2 \times \frac{3}{6} \\
&\quad + (4-3.5)^2 \times \frac{4}{6} + (5-3.5)^2 \times \frac{5}{6} + (6-3.5)^2 \times \frac{6}{6} \\
&= (2.5)^2 \times \frac{1}{3} + (1.5)^2 \times \frac{1}{3} + (0.5)^2 \times \frac{1}{3} \\
&= \frac{8.75}{3} \fallingdotseq 2.92
\end{aligned}
$$

また, $\sigma \fallingdotseq \sqrt{2.92} \fallingdotseq 1.71$ となります.

今度は期待値のところで出てきた次の確率密度関数

$$
f(x) = \begin{cases} xe^{x} & (x \geq 0) \\ 0 & (x < 0) \end{cases}
$$

この確率分布の分散を求めると, $\mu = 2$ なので

$$
\sigma^2 = \int_0^\infty (x-2)^2 x e^{-x} dx = \int_0^\infty (x^3 - 4x^2 + 4x) e^{-x} dx
$$

ここで, ガンマ関数

$$
\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt
$$

を使うと

$$
\sigma^2 = \Gamma(4) - 4\Gamma(3) + 4\Gamma(2) = 3! - 4 \cdot 2! + 4 \cdot 1! = 2
$$

また, $\sigma = \sqrt{2} \fallingdotseq 1.41$ です.

それでは, 一般的な指数分布の期待値と分散も求めてみます.指数分布の確率密度関数は

$$
f(x) = \begin{cases} \lambda e^{-\lambda x} & (x \geq 0) \\ 0 & (x < 0) \end{cases}
$$

です.まず期待値は

$$
\mu = \int_{-\infty}^{\infty} x f(x) dx = \int_0^\infty x \lambda e^{-\lambda x} dx
$$

ですので, 部分積分を利用すると

$$
\int_0^\infty x \lambda e^{-\lambda x} dx = \left[-\frac{x}{\exp(\lambda x)} \right]_0^\infty + \int_0^\infty e^{-\lambda x} dx
$$

となり, この右辺の第 1 項の中の極限は

$$
\lim_{x \to \infty} \frac{x}{\exp(\lambda x)} = \frac{\infty}{\infty}
$$

となって不定形となるので, この場合, 分母と分子それぞれの極限を求めて

$$\lim_{x\to\infty}\frac{x}{\exp(\lambda x)} = \lim_{x\to\infty}\frac{(x)'}{\{\exp(\lambda x)\}'} = \lim_{x\to\infty}\frac{1}{\lambda\exp(\lambda x)} = 0$$

よって

$$\mu = \int_0^\infty x\lambda e^{-\lambda x}dx = \int_0^\infty e^{-\lambda x}dx = \left[-\frac{\exp(-\lambda x)}{\lambda}\right]_0^\infty = \frac{1}{\lambda}$$

次に x^2 の期待値を求めます.

$$E[x^2] = \int_{-\infty}^\infty x^2 f(x)dx = \int_0^\infty x^2\lambda\exp(-\lambda x)dx$$

ですから, 部分積分を使って

$$\begin{aligned}\int_0^\infty x^2\lambda\exp(-\lambda x)dx &= \left[-\frac{x^2}{\exp(\lambda x)}\right]_0^\infty + \int_0^\infty x^2\lambda\exp(-\lambda x)dx \\ &= 0 + \int_0^\infty x^2\lambda\exp(-\lambda x)dx \\ &= \int_0^\infty x^2\lambda\exp(-\lambda x)dx\end{aligned}$$

もう一度, 部分積分を使うと

$$\begin{aligned}\int_0^\infty x^2\lambda\exp(-\lambda x)dx &= \left[-\frac{2x}{\lambda\exp(-\lambda x)}\right]_0^\infty + 2\int_0^\infty\frac{\exp(-\lambda x)}{\lambda}dx \\ &= 2\left[-\frac{\exp(-\lambda x)}{\lambda^2}\right]_0^\infty = \frac{2}{\lambda^2}\end{aligned}$$

よって, 分散は

$$V[X] = E[x^2] - \mu^2 = \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2}$$

となります.

14. チェビシェフの不等式

確率変数と確率分布から, 期待値と分散を求めることができました.ここで, 確率分布が分からない場合でも, 確率変数から平均値と分散を求めることができます.期待値は平均値と同じです.このとき, ある確率変数の確率について, どれくらいか見当をつけられないでしょうか.例えば, 期末テストの数学の結果から平均と分散がわかったとして, 平均点の ± 10 に入る確率はどのくらいになるかを考えてみます.
まず, 分散を求める式は

$$\sigma^2 = \int_{-\infty}^\infty (X-\mu)^2 f(x)dx$$

$(X-\mu)^2$ は平方なので常に正の値, そして確率密度関数 $f(x)$ も $f(x) \geq 0$ なので, この積分は常に正の値となります.ここで, $\mu - k\sigma$ から $\mu + k\sigma$ までの積分の値を取り除きます.つまり, 積分領域を 3 つに分けて

$$\sigma^2 = \int_{-\infty}^{\mu-k\sigma} (X-\mu)^2 f(x)dx + \int_{\mu-k\sigma}^{\mu+k\sigma} (X-\mu)^2 f(x)dx + \int_{\mu+k\sigma}^{\infty} (X-\mu)^2 f(x)dx$$

この式から第 2 項の積分を取り除くと, 元の値以下となるので不等号がつきます

$$\sigma^2 \geq \int_{-\infty}^{\mu-k\sigma} (X-\mu)^2 f(x)dx + \int_{\mu+k\sigma}^{\infty} (X-\mu)^2 f(x)dx$$

次に, $(X-\mu)^2$ の X に $\mu - k\sigma$ と $\mu + k\sigma$ を代入すると

$$(\mu - k\sigma - \mu)^2 = k^2\sigma^2, \quad (\mu + k\sigma - \mu)^2 = k^2\sigma^2$$

よって

$$(X-\mu)^2 = k^2\sigma^2$$

になります.これを

$$\sigma^2 \geq \int_{-\infty}^{\mu-k\sigma} (X-\mu)^2 f(x)dx + \int_{\mu+k\sigma}^{\infty} (X-\mu)^2 f(x)dx$$

に代入すると, $k^2\sigma^2$ が積分の外に出せるので

$$\sigma^2 \geq k^2\sigma^2 \int_{-\infty}^{\mu-k\sigma} f(x)dx + k^2\sigma^2 \int_{\mu+k\sigma}^{\infty} f(x)dx$$

となります.この右辺の第 1 項と第 2 項の積分はそれぞれ

$$P(X < \mu - k\sigma) = \int_{-\infty}^{\mu-k\sigma} f(x)dx$$
$$P(X > \mu + k\sigma) = \int_{\mu-k\sigma}^{\infty} f(x)dx$$

と表せるので, 絶対値記号を使ってまとめると

$$\sigma^2 \geq k^2\sigma^2 P(|X-\mu| \geq k\sigma)$$

そして, この両辺を $k^2\sigma^2$ で割ると

$$\frac{1}{k^2} \geq P(|X-\mu| \geq k\sigma) \quad (k > 0)$$

となります.これをチェビシェフの不等式といいます.この不等式は, ある確率変数 X が平均値 μ から標準偏差 σ の k 倍以上離れている確率は全体の $\frac{1}{k^2}$ より小さいということを表しています.これを図 7 を見てみると, 確率は色が付いた部分の面積に相当します.例えば, $2k$ 以上なら $1/4$ 以下, $3k$ 以上なら $1/9$ 以下になります.これは, 確率分布がどのようなものでも成り立ちます.

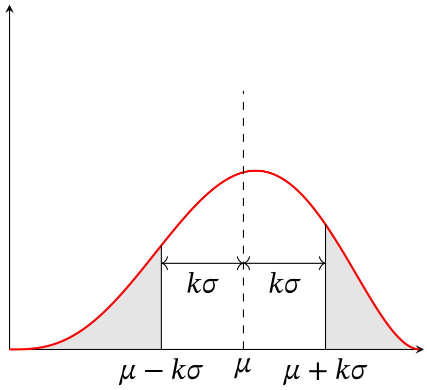


Fig.7: チェビシェフの不等式

それでは, チェビシェフの不等式を使って, 最初の問題を考えてみます.期末テストの数学の結果から平均と分散がわかったとして, 平均点の ±10 に入る確率はどのくらいになるかでした.ここで, 平均は 60, 分散は 484, 標準偏差は 22 とすると, 平均点の ±10 は 50, 70 になるので, それを満たす k は 11/5 となります.よって

$$P(|X-\mu| \leq \frac{11}{5}\sigma) \leq \frac{11}{5}^2 = 0.44$$

となって, 50 点から 70 点になる確率は 44% 以下であるということがわかります.この平均と分散のヒストグラムは次のようになっていました.

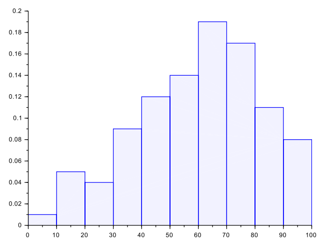


Fig.8: 平均点 60, 標準偏差 22 のヒストグラム

50 点から 70 点までの範囲を見てみると 44% 以下になっていることがわかります.このように, 分布がわからなくても, 平均と分散がわかっていれば, ある確率変数がどれくらいなのかがある程度当たりをつけることができます.

15. 2 項分布

ある試行が別の試行に影響を及ぼさないことを独立試行, またはベルヌーイ試行といいました.独立試行のところで説明しましたが, もう一度簡単に見ていきます.ここで, サイコロを振ったときに 1 の目が出る確率は 1/6 で, それ以外は 5/6 となります.サイコロを 2 回振ったときに 1 の目が出る確率は独立試行なので,

$$\frac{1}{6} \times \frac{1}{6} = \frac{1}{36}$$

となります.それでは, 5 回振ったときに 1 の目が出る回数を確率変数 X としたとき, 1 の目が出る回数が 2 回, つまり $X(2)$ の確率を考えてみます.まず, 5 回サイコロを振ったときに 1 の目が出る組合せは, 5 個の異なるものから 2 個を取り出すことなので

$$
{}_5C_2 = \frac{5 \cdot 4}{2} = 10
$$

となり，10 通りです.次に，1 の目が出る回数が 2 回，それ以外が 3 回なので，この確率は

$$
\left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right)^3
$$

となります.よって，$X(2)$ となる確率は

$$
P(X = 2) = 10 \times \left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right)^3 = 10 \times \frac{1}{36}\frac{125}{216} = \frac{625}{3888}
$$

となります.一般に，ある事象 A の起こる確率 $P(A) = p$ が与えられているとき，n 回独立試行を行って，A が x 回起こる確率は

$$
f(x) = {}_nC_r p^n (1-p)^{n-x}
$$

となります.このような確率分布を **2 項分布**または**ベルヌーイ分布**といいます. 2 項分布は n 回の独立試行回数と，事象 A の起こる確率を p としたとき，$Bin(n, p)$ と表します.サイコロの例だと 5 回振って 1 の目が出る確率が $1/6$ なので，2 項分布 $Bin(5, 1/6)$ に従います. 2 項分布で n を固定したものと p を固定したものを図 9 に示しました.

(a) 2 項分布 $Bin(20, p)$

(b) 2 項分布 $Bin(n, 1/6)$

Fig.9: 2 項分布

ここで，2 項分布と 2 項定理の関係を見てみましょう.まず，2 項分布は

$$
f(x) = {}_nC_x p^n (1-p)^{n-x}
$$

でした. $q = 1 - p$ とおくと

$$
f(x) = {}_nC_x p^n q^{n-x}
$$

と変形できます. 2 項定理は

$$
(a + b)^n = \sum_{x=0}^{n} {}_nC_x a^{n-x} b^x
$$

です.ここで

$$
\sum_{x=0}^{n} f(x) = f(0) + f(1) + \cdots + f(n)
$$

という式を考えると, この右辺は

$$ {}_nC_0 p^0 q^{n-0} + {}_nC_1 p^1 q^{n-1} + \cdots + {}_nC_n p^n q^{n-n} $$

となって, 2 項定理の式と一致するので

$$ \sum_{x=0}^{n} f(x) = (p+q)^n $$

という関係がわかります. $p+q=1$ なので

$$ \sum_{x=0}^{n} f(x) = 1 $$

となります.これは $f(x)$ が確率密度関数の性質を持っているということです.

それでは, 2 項分布の平均と分散を求めてみます. 2 項分布の確率変数は離散なので, 平均は

$$ \mu = \sum_{x=0}^{n} x f(x) $$

$f(x)$ を置き換えると

$$ \mu = \sum_{x=0}^{n} x {}_nC_x p^x q^{n-x} $$

これを変形すると

$$ \mu = \sum_{x=0}^{n} x \frac{n!}{x!(n-x)!} p^x q^{n-x} $$

となります. $x=0$ のとき, この項は 0 になるので

$$ \mu = \sum_{x=1}^{n} x \frac{n!}{x!(n-x)!} p^x q^{n-x} $$

分母の $x!$ を $x(x-1)!$ とすると,

$$ \mu = \sum_{x=1}^{n} x \frac{n!}{x(x-1)!(n-x)!} p^x q^{n-x} = \sum_{x=1}^{n} \frac{n!}{(x-1)!(n-x)!} p^x q^{n-x} $$

$n! = n(n-1), \quad p^x = p \cdot p^{x-1}, \quad n-x = (n-1)-(x-1)$ と変形すると

$$ \mu = \sum_{x=1}^{n} np \frac{(n-1)!}{(x-1)!\{(n-1)-(x-1)\}!} p^{x-1} q^{(n-1)-(x-1)} $$

ここで,

$$ t = x-1 \qquad m = n-1 $$

とおくと

$$\mu = \sum_{t=0}^{m} np \frac{m!}{t!(m-t)!} p^t q^{t-m} = np \sum_{t=0}^{m} {}_mC_t p^t q^{t-m}$$

np の後ろの部分が 2 項定理のかたちをしているので

$$\mu = np(p+q)^m$$

$p+q=1$ だから

$$\mu = np$$

となります.続いて分散は

$$\sigma^2 = E[X^2] - \mu^2$$

なので,

$$E[X^2] = \sum_{x=0}^{n} x^2 f(x) = \sum_{x=0}^{n} x^2 {}_nC_x p^x q^{n-x}$$

平均と同じように変形して

$$t = x-1 \qquad m = n-1 \qquad \therefore x = t+1$$

とおくと

$$\begin{aligned}
E[X^2] &= np \sum_{t=0}^{m} (t+1) \frac{m!}{t!(m-t)!} p^t q^{t-m} \\
&= np \sum_{t=0}^{m} t \frac{m!}{t!(m-t)!} p^t q^{t-m} + np \sum_{t=0}^{m} \frac{m!}{t!(m-t)!} p^t q^{t-m} \\
&= np(mp) + np \cdot 1 \\
&= n(n-1)p^2 + np
\end{aligned}$$

よって

$$\sigma^2 = E[X^2] - \mu^2 = n(n-1)p^2 + np - (np)^2 = np(1-p)$$

$$\therefore \sigma^2 = np(1-p)$$

2 項分布の平均と分散が求まりました.ここで, $X/n = t$ という新しい確率変数を考えます. $f(x)$ が 2 項分布 $Bin(n, 0.2)$ に従うとき, t の分布の分布は図 10 のようになります.

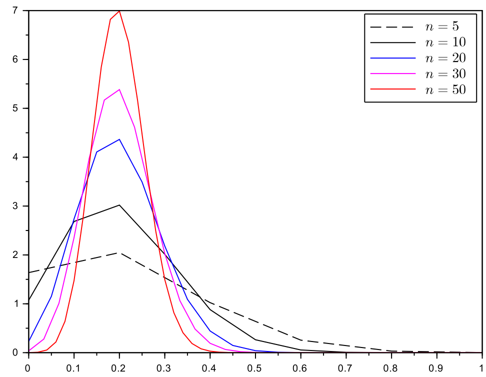


Fig.10: $t = X/n$ の分布

この図から n を大きくしていくと, $X/n = 0.2$ のまわりに分布が集中していくことがわかります. X の平均は np なので, X/n の平均は p となって, n によらずに一定になります.また, X の分散は $np(1-p)$ なので, X/n の分散は $p(1-p)/n$ となって, n が大きくなると p が 0 に収束していき, 中心付近に集中します.ここで, チェビシェフの不等式

$$\frac{1}{k^2} \geq P(|X-\mu| \geq k\sigma) \quad (k>0)$$

を利用して, $\mu = np, \quad \sigma = \sqrt{np(1-p)}$ とすると, 任意の正の数 k に対して

$$P(|X-np| \leq k\sqrt{np(1-p)}) \geq 1 - \frac{1}{k^2}$$

が成り立ちます.確率 P は 1 を超えないので

$$1 \geq P(|X-np| \leq k\sqrt{np(1-p)}) \geq 1 - \frac{1}{k^2}$$

カッコ内の両辺を n で割ると

$$1 \geq P\left(|t-p| \leq k\sqrt{\frac{p(1-p)}{n}}\right) \geq 1 - \frac{1}{k^2}$$

この式で k を大きくしても, $\sqrt{n}$ をそれよりも大きくすれば, $k\sqrt{\frac{p(1-p)}{n}}$ はいくらでも小さくすることができます.ここで

$$\epsilon = k\sqrt{\frac{p(1-p)}{n}}$$

とおくと, 次のように変形できます

$$1 \geq P(p-\epsilon \leq t \leq p+\epsilon) \geq 1 - \frac{1}{k^2}$$

k を十分大きくすれば, 右辺はほとんど 1 に等しくなり, n を k^2 よりも十分大きくすれば, ϵ は非常に小さくなって, t が p に近づいていきます.これは, 試行回数 n を増やすほど, その事象の起こる割合は一定の値 p に近づくということを意味しています.この性質を大数の法則といいます.チェビシェフの不等式はどの確率分布でも成り立ち, 大数の法則は統計的確率の

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

の根拠となっています.

16. 多項分布

サイコロを振ったときに 1 の目が出る確率と 1 以外の目が出る確率は 2 項分布に従います.では，5 回サイコロを振った場合に 1 の目がでる回数を 1 回，2 の目が出る回数を 2 回，3 の目がでる回数を 2 回というように，3 つ以上の結果が起こる確率を求めます.各目が出る確率を $p_1, p_2, p_3, p_4, p_5, p_6$ とすると, どの目も 1/6 であり,

$$p_1 + p_2 + p_3 + p_4 + p_5 + p_6 = 1$$

となります.この場合，$(p_1 + p_2 + p_3 + p_4 + p_5 + p_6)^5$ で，p_1 を 1 個，p_2 を 2 個，p_3 を 2 個選ぶ組合せの数で確率が決まるので多項定理

$$(a_1 + a_2 + \cdots + a_n)^n = \sum \frac{n!}{n_1!n_2!\dots n_m!} a_1^{n_1} a_2^{n_2} \dots a_m^{n_m}$$

$$(n_1 + n_2 + \cdots + n_m = n)$$

を利用して，$n_1 = 1, n_2 = 2, n_3 = 2, n_4 = n_5 = n_6 = 0, n = 5$ とすると

$$\frac{5!}{1!2!2!0!0!0!} \left(\frac{1}{6}\right)^1 \left(\frac{1}{6}\right)^2 \left(\frac{1}{6}\right)^2 \left(\frac{1}{6}\right)^0 \left(\frac{1}{6}\right)^0 \left(\frac{1}{6}\right)^0 \fallingdotseq 0.0038$$

となります.一般に，1 回の試行で起こりえる結果が m 通りあり, それぞれの起きる確率を $p_1, p_2, \dots, p_m$ とし

$$p = p_1 + p_2 + \cdots + p_m = 1$$

という関係であるとき，n 回の独立試行を行い，i 番目の結果の起こる回数を確率変数 X_i としたとき

$$X_1 = n_1, X_2 = n_2, \dots, X_m = n_m$$

となる確率は, 確率密度関数

$$f(n_1, n_2, \dots, n_m) = \frac{n!}{n_1!n_2!\cdots n_m!} p_1^{n_1} p_2^{n_2} \cdots p_m^{n_m}$$

に従います.この確率分布を多項分布といいます.また, 確率変数 X_i はそれぞれ 2 項分布なので,

$$\mu_i = np_i, \quad \sigma_i^2 = np_i(1 - p_i)$$

となります.

17. ポアソン分布

あるゲームにおいて, 宝箱から貴重品を手に入れる確率が $1/100$ だとします.貴重品が手に入る数を確率変数とすると, これは 2 項分布に従いますから, 手に入る貴重品の数は

$$p(X=x)=f(x)={}_nC_x p^x(1-p)^{n-x}$$

となります.よって, 貴重品が手に入らない確率は

$$f(0)={}_{100}C_0\left(\frac{1}{100}\right)^0\left(1-\frac{1}{100}\right)^{100}=\left(\frac{99}{100}\right)^{100}\fallingdotseq 0.366$$

貴重品が 1 個手に入る確率は

$$f(1)={}_{100}C_1\left(\frac{1}{100}\right)^1\left(1-\frac{1}{100}\right)^{99}=\left(\frac{99}{100}\right)^{99}\fallingdotseq 0.370$$

同様に

$$f(2)\fallingdotseq 0.185,\quad f(3)\fallingdotseq 0.061,\quad f(4)\fallingdotseq 0.015,\quad f(5)\fallingdotseq 0.003$$

となって, x が $3,4,5,\ldots$ と大きくなると, その確率は急速に 0 に近づいていきます.このようにめったに起こらない事象に対して, 何回も試行を行うときには, 2 項分布を近似した分布を考えることができます. 2 項分布

$$f(x)={}_nC_x p^x(1-p)^{n-x}=\frac{n!}{x!(n-x)!}p^x(1-p)^{n-x}$$

を展開して

$$f(x)=\frac{n!}{x!(n-x)!}p^x(1-p)^{n-x}=\frac{n(n-1)(n-2)\cdots(n-x+1)}{x!}p^x(1-p)^{n-x}$$

$\mu=np$ だから $p=\mu/n$ を代入して

$$f(x)=\frac{n(n-1)(n-2)\cdots(n-x+1)}{x!}\left(\frac{\mu}{n}\right)^x\left(1-\frac{\mu}{n}\right)^{n-x}$$

次のように変形して

$$f(x)=\frac{\mu^x}{x!}\cdot 1\cdot\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{x-1}{n}\right)\left(1-\frac{\mu}{n}\right)^{n-x}$$

最後の項は 2 つの式に分けて

$$f(x)=\frac{\mu^x}{x!}\cdot 1\cdot\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{x-1}{n}\right)\left(1-\frac{\mu}{n}\right)^{-x}\left(1-\frac{\mu}{n}\right)^{n}$$

ここで $n\to\infty$ の極限を考えると,

$$\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{x-1}{n}\right),\quad\left(1-\frac{\mu}{n}\right)^{-x}$$

はすべて 1 になります.最後の項は

$$\lim_{n\to\infty}\left(1-\frac{\mu}{n}\right)^n = \lim_{n\to\infty}\left\{\left(1-\frac{\mu}{n}\right)^{-\frac{n}{\mu}}\right\}^{-\mu}$$

と変形し，$p=\mu/n$ だから

$$\lim_{n\to\infty}\left(1-\frac{\mu}{n}\right)^n = \lim_{n\to\infty}\left\{(1-p)^{-\frac{1}{p}}\right\}^{-\mu}$$

$p=-1/m$ とおくと

$$\lim_{n\to\infty}\left(1-\frac{\mu}{n}\right)^n = \lim_{n\to\infty}\left\{\left(1-\frac{1}{m}\right)^{m}\right\}^{-\mu}$$

ここで e の公式

$$e=\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n$$

から

$$\lim_{n\to\infty}\left(1-\frac{\mu}{n}\right)^n = e^{-\mu}$$

となって, 結局

$$f(x)=\frac{\mu^x}{x!}e^{-\mu}$$

となります.この分布をポアソン分布といい, 上記の式はポアソン分布の確率密度関数です.ポアソン分布は平均 $\mu=np$ の 2 項分布において，$n\to\infty$ の極限をとっているので, 試行回数 n が大きいときに対応し，p の極限は 0 になります.平均は $\mu=np$ でポアソン分布は $P(\mu)$ と表します.ポアソン分布のグラフを図 11 に示します.

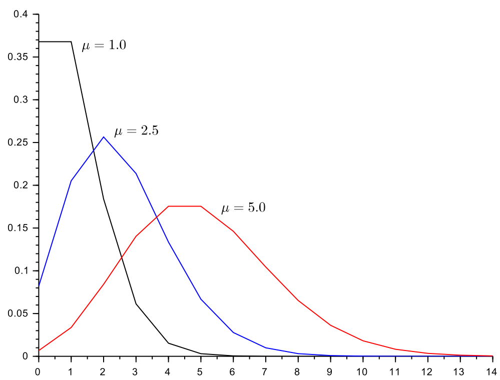


Fig.11: ポアソン分布

分散は 2 項分布の分散 $\sigma^2=np(1-p)$ の極限なので，$p=\mu/n$ を代入して

$$\sigma^2=\lim_{n\to\infty}np(1-p)=\lim_{n\to\infty}\mu\left(1-\frac{\mu}{n}\right)=\mu$$

つまり, ポアソン分布の分散 σ^2 は平均 μ と等しくなります.

18. 超幾何分布

2 項分布や多項分布, ポアソン分布では独立試行のため, 複数の試行を行っても, その結果が他の試行に影響することはありません. では, 試行すると確率が変わる場合はどうなるでしょうか. 次のような場合を考えてみます. 箱の中に赤い玉が 2 個, 白い玉が 3 個入っています. このとき, 3 個の玉を箱から取り出したときに, 赤い玉が 2 個, 白い玉が 1 個になる確率を求めます. 条件を満たす組合せは「赤赤白」「赤白赤」「白赤赤」の 3 通りです. すべての場合の数は ${}_5C_3$ なので

$$P = \frac{3}{{}_5C_3} = \frac{3}{10}$$

となります. これは, 赤い玉 2 個から 2 個を選ぶ場合の数 ${}_2C_2$ と, 白い玉 3 個から 1 個を選ぶ場合の数 ${}_3C_1$ の積が起こり得る場合の数なので

$${}_2C_2 \times {}_3C_1 = 3$$

と計算することができます. これを一般に考えると, 赤い玉が m 個, 白い玉が n 個入った箱があるとします. この箱から玉を n 個取り出したときに, 赤い玉が k 個含まれる確率を求めると, すべての事象の数は $m+n$ 個から r 個を選ぶ場合の数なので

$${}_{m+n}C_r$$

次に, 確率を求める事象は, 赤い玉 m 個の中から k 個を, 白い玉 n 個から $r-k$ 個を選ぶ場合の数なので

$${}_mC_k \times {}_nC_{r-k}$$

となって, 赤い玉を取り出す数を確率変数 $X = x$ とすれば

$$f(x) = \frac{{}_mC_x \times {}_nC_{r-x}}{{}_{m+n}C_r}$$

で与えられます. k は 0 から r まで変化し

$$\sum_{x=0}^{r} f(x) = 1$$

という性質を持っているので, これは確率密度関数です. この確率分布を超幾何分布といいます.

これまでの場合は, 赤い玉と白い玉の 2 種類でしたが, それ以上の場合を考えることもできます. いま, 箱の中に赤い玉が a 個, 白い玉が b 個, 青い玉が c 個入っているとします. その中から n 個の玉を取り出したときに, 赤い玉が n_1 個, 白い玉が n_2 個, 青い玉が n_3 個含まれる確率を求めます. まず,

$$n = n_1 + n_2 + n_3$$

を満たす必要があります. そして, すべての場合の数は $a+b+c$ 個から n 個を選ぶので

$${}_{a+b+c}C_n$$

となります. つぎに, 赤い玉 n_1 個, 白い玉 n_2 個, 青い玉 n_3 個含まれる場合の数は

$${}_aC_{n_1} \times {}_bC_{n_2} \times {}_cC_{n_3}$$

となって, 求める確率は

$$\frac{{}_aC_{n_1} \times {}_bC_{n_2} \times {}_cC_{n_3}}{{}_{a+b+c}C_n}$$

となります.これは玉の種類が増えても同様に計算することができます.

19. 幾何分布

次のような場合を考えます.サイコロを 3 回振ったときに, 1 の目が 1 回でも出てくる確率はどれくらいでしょうか.これは

- A: 1 回目に 1 の目が出る
- B: 1 回目は 1 以外の目で, 2 回目に 1 が出る
- C: 1 回目と 2 回目は 1 以外の目で, 3 回目に 1 が出る

それぞれの確率を計算すると

$$P(A) = \frac{1}{6}, \quad P(B) = \frac{5}{6} \times \frac{1}{6} = \frac{5}{36}, \quad P(C) = \left(\frac{5}{6}\right)^2 \times \frac{1}{6} = \frac{25}{216}$$

求める確率はこれらの和の事象なので

$$\frac{1}{6} + \frac{5}{36} + \frac{25}{216} = \frac{36 + 30 + 25}{216} = \frac{91}{216} \doteqdot 0.42$$

となります.もし, 4 回目に初めて 1 が出る確率を考えると

$$\left(\frac{5}{6}\right)^3 \times \frac{1}{6}$$

となることがわかります.同様に r 回目に初めて 1 が出る確率は

$$\left(\frac{5}{6}\right)^{r-1} \times \frac{1}{6}$$

となります.一般に, 独立試行を行ったときに事象 A の起こる確率を p とすると, 事象 A が最初に発生するまでの回数 $x = 1, 2, 3, \ldots$ を確率変数とすれば

$$f(x) = p(1-p)^{x-1}$$

で得られます.これは試行を繰り返せばいずれ事象 A が起きることになるので,

$$\sum_{x=1}^{\infty} f(x) = 1$$

の性質を持っていますから, 確率密度関数です.この確率分布を幾何分布といいます.この確率密度関数は初項が p で, 公比が $1-p$ の等比数列の形をしています.幾何分布は, 時間を $1, 2, 3, \cdots$ と離散的に考えるとき, 初めて A が起きるまで待つ時間の確率分布と考えることができ, 待ち時間分布とも呼ばれます.

次に幾何分布の平均と分散を求めてみます.平均は

$$\mu = \sum_{x=1}^{\infty} x f(x) = \sum_{x=1}^{\infty} x p(1-p)^{x-1} = p \sum_{x=1}^{\infty} x(1-p)^{x-1}$$

ここで $1/(1-x)$ のマクローリン展開を考えると

$$\frac{1}{1-x} = 1 + x + x^2 + \cdots = \sum_{k=0}^{\infty} x^k$$

この両辺を微分すると

$$\frac{1}{(1-x)^2} = \sum_{k=1}^{\infty} k x^{k-1}$$

$x = 1-p, \quad k = x$ とおくと

$$\frac{1}{(1-(1-p))^2} = \sum_{x=1}^{\infty} x(1-p)^{x-1}$$

この右辺は幾何分布の平均の式

$$\mu = p \sum_{x=1}^{\infty} x(1-p)^{x-1}$$

の数列の部分と一致するので

$$\mu = p \times \frac{1}{(1-(1-p))^2} = \frac{p}{p^2} = \frac{1}{p}$$

別の計算方法として，無限和を求めるやり方があります．平均の式は

$$\mu = \sum_{x=1}^{\infty} x p(1-p)^{x-1}$$

右辺は，$(1-p)$ を掛けたものを引くと，係数 x が消え，それは確率密度関数と同じなので 1 になります．

$$\sum_{x=1}^{\infty} x p(1-p)^{x-1} - (1-p) \sum_{x=1}^{\infty} x p(1-p)^{x-1} = 1$$

整理すると

$$\begin{aligned} (1-(1-p)) \sum_{x=1}^{\infty} x p(1-p)^{x-1} &= 1 \\ \sum_{x=1}^{\infty} x p(1-p)^{x-1} &= \frac{1}{p} \end{aligned}$$

次に分散を求めます．この分散は

$$\sigma^2 = E[X^2] - \mu^2$$

の式を使います．また

$$E[X^2] = E[X(X-1)] + E[X]$$

という関係を利用するので，$E[X(X-1)]$ をまず求めます.そのために, 平均を求めたときに出てきた次の式を使います.

$$\frac{1}{(1-x)^2} = \sum_{k=1}^{\infty} kx^{k-1}$$

この両辺をもう一度微分します.

$$\frac{2}{(1-x)^3} = \sum_{k=1}^{\infty} k(k-1)x^{k-2}$$

この両辺に x をかけて，$x = 1-p, \quad k = x$ とおくと

$$\begin{aligned}
\frac{2x}{(1-x)^3} &= \sum_{k=1}^{\infty} k(k-1)x^{k-1} \\
\frac{2(1-p)}{(1-(1-p))^3} &= \sum_{x=1}^{\infty} x(x-1)(1-p)^{x-1} \\
\frac{2(1-p)}{p^3} &= \sum_{x=1}^{\infty} x(x-1)(1-p)^{x-1}
\end{aligned}$$

よって

$$E[X(X-1)] = p\sum_{x=1}^{\infty} x(x-1)(1-p)^{x-1} = \frac{2(1-p)}{p^2}$$

結局, 分散は

$$\begin{aligned}
\sigma^2 &= E[X^2] - \mu^2 \\
&= E[X(X-1)] + E[X] - \mu^2 \\
&= \frac{2(1-p)}{p^2} + \frac{1}{p} + \left(\frac{1}{p}\right)^2 \\
&= \frac{2(1-p) - 1 + p}{p^2} \\
&= \frac{1-p}{p^2}
\end{aligned}$$

幾何分布を使った例を見る前に, この累積分布関数を求めます.幾何分布は離散の確率分布なので, 累積密度関数は

$$F(x) = P(X \leq x) = \sum_{k=1}^{x} f(x) = \sum_{k=1}^{k} p(1-p)^{k-1} = p\sum_{k=1}^{x}(1-p)^{k-1}$$

等比数列の和 S は, 初項 a, 公比 r, 項数 n とすると

$$S = \frac{a(r^2-1)}{r-1} = \frac{a(1-r^n)}{1-r}$$

計算すると

$$\sum_{k=1}^{x}(1-p)^{k-1} = \frac{1-(1-p)^x}{p}$$

よって

$$F(x) = p\sum_{k=1}^{x}(1-p)^{k-1} = p\frac{1-(1-p)^x}{p} = 1-(1-p)^x$$

となります.それでは,累積分布関数を使って最初のサイコロの例の確率を計算してみると

$$F(x) = 1-(1-p)^x = 1-\left(1-\frac{1}{6}\right)^3 = 1-\frac{125}{216} = \frac{91}{216} \fallingdotseq 0.42$$

となって一致します.今度は別の事例を見てみましょう.あるスマホゲームでガチャがあります.このガチャで SSR カードが出る確率はどのようになっているか,ガチャを回した回数との関係を図 12 に示します.

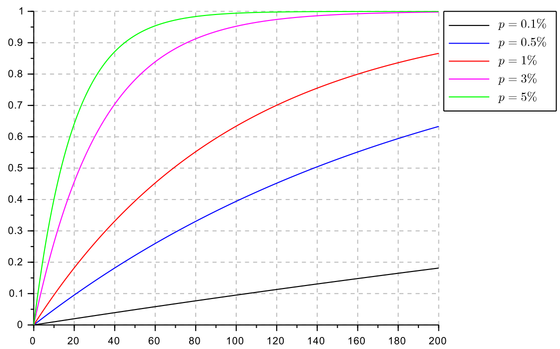


Fig.12: 幾何分布

横軸がガチャを回した回数,縦軸が SSR カードを最初に引く確率, p が SSR カードの出る確率です.

20. ガウス積分

正規分布(ガウス分布)をはじめる前に,ガウス積分について見ていきます.ガウス積分は以下の式

$$f(x) = e^{-x^2}$$

の形をした関数を $-\infty$ から ∞ まで積分した値です.ここで,この積分の値を I とすると

$$I = \int_{-\infty}^{\infty} e^{-x^2}dx$$

これとまったく同じ式で x を y に置き換えたものを考え

$$I = \int_{-\infty}^{\infty} e^{-y^2}dy$$

これらの積を求めると

$$I^2 = \int_{-\infty}^{\infty} e^{-x^2}dx \cdot \int_{-\infty}^{\infty} e^{-y^2}dy$$

まとめると

$$I^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp(-(x^2+y^2))dxdy$$

重積分の形になります.この積分は図 13 に示すように

$$z = exp(-(x^2+y^2))$$

という関数に体積になっています.

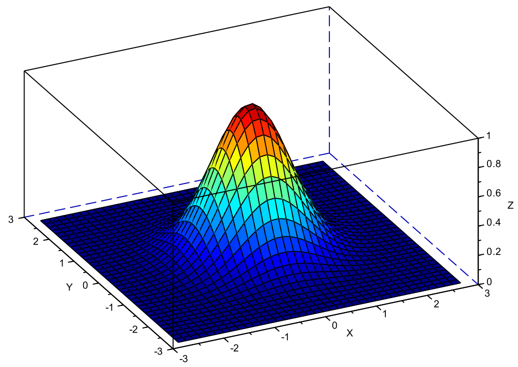


Fig.13: $z = exp(-(x^2+y^2)$

この重積分を計算するために,直交座標(x,y)から極座標(r,θ)に変数変換します.すると

$$x^2+y^2=r^2$$

となって,微分係数は

$$dxdy \rightarrow rdrd\theta$$

となります.これは直交座標での微小面積 $dxdy$ が,極座標での微小面積 $(rdrd\theta)$ に変換されています.また,積分範囲は

$$-\infty \le x \le \infty, \quad -\infty \le y \le \infty \quad \rightarrow \quad 0 \le r \le \infty, \quad 0 \le \theta \le 2\pi$$

に変わります.よって

$$I^2 = \int_0^{2\pi}\int_0^{\infty} \exp(-r^2)rdrd\theta$$

となります.まず

$$\int_0^{\infty} \exp(-r^2)rdr$$

の積分を求めるには,置換積分を使って $t=r^2, \quad dr = dt/2r$ とすると

$$\int_0^{\infty} \exp(-r^2)rdr = \int_0^{\infty} \frac{1}{2}\exp(-t)dt = \left[-\frac{1}{2}\exp(-t)\right]_0^{\infty} = \frac{1}{2}$$

よって

$$I^2 = \int_0^{2\pi} \frac{1}{2} d\theta = \left[\frac{\theta}{2} \right]_0^{2\pi} = \pi$$

$$\therefore I = \pm\sqrt{\pi}$$

ただし，I の値は正なので

$$\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}$$

となります．

以降はガウス積分の類似した形のものをいくつか見ていきます．まずは

$$\int_{-\infty}^{\infty} e^{-ax^2} dx \quad (a > 0)$$

これは $t = \sqrt{a}x, \quad dt = \sqrt{a}dx$ とおくと

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

右辺の積分はガウス積分なので $\sqrt{\pi}$ となるから

$$\int_{-\infty}^{\infty} e^{-ax^2} dx = \frac{\sqrt{\pi}}{\sqrt{a}} = \sqrt{\frac{\pi}{a}}$$

となります．次に

$$\int_0^{\infty} e^{-ax^2} dx \quad (a > 0)$$

これは e^{-ax^2} のグラフ(図 14)を見てください．

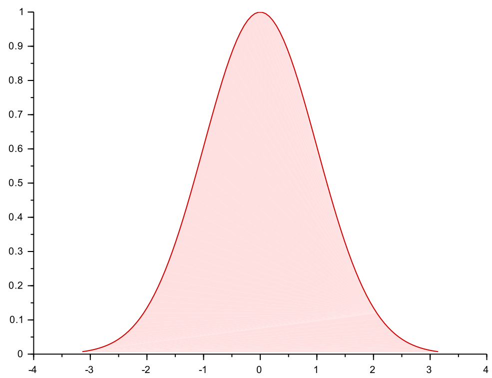


Fig.14: e^{-ax^2}

y 軸に対して $x = 0$ のところで軸対称になっていることがわかります.このような性質をもつ関数を偶関数といいます.つまり, e^{-ax^2} は偶関数です.ここで積分範囲の下限が $-\infty$ から 0 に変わったということは,グラフ上で色が付いている面積の半分ということになります.よって

$$\int_0^{\infty} e^{-ax^2} dx = \frac{1}{2}\sqrt{\frac{\pi}{a}}$$

となります.次は

$$\int_0^{\infty} xe^{-ax^2} dx \quad (a > 0)$$

そのまま,積分を計算すると

$$\int_0^{\infty} xe^{-ax^2} dx = \left[-\frac{1}{2a}e^{-ax^2}\right]_0^{\infty} = -\left(-\frac{1}{2a}\right) = \frac{1}{2a}$$

となります.最後に

$$\int_{-\infty}^{\infty} x^{2n}e^{-ax^2} dx, \qquad \int_0^{\infty} x^{2n+1}e^{-ax^2} dx \quad (a > 0)$$

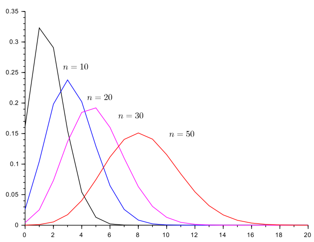
それぞれ

$$\int_{-\infty}^{\infty} x^{2n}e^{-ax^2} dx = \frac{n!}{2a^{n+1}}$$
$$\int_{-\infty}^{\infty} x^{2n+1}e^{-ax^2} dx = \frac{(2n-1)!!}{2^n a^n}\sqrt{\frac{\pi}{a}}$$

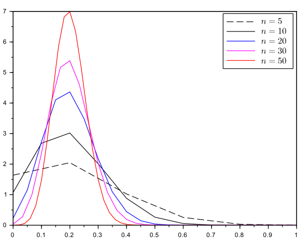
となります.

21. 正規分布

2 項分布で n を大きくしていくと,分布が特定の形になっていくことが知られています.また, 2 項分布のところで $t = X/n$ の分布は n を大きくしていくと,大数の法則によって, p のまわりに集中していくことがわかりました.それぞれ図で確認してみると



(a) 2 項分布 $Bin(n, 1/6)$



(b) $t = X/n$ の分布

Fig.15: 2 項分布

このように n を大きくしていくと,その分布は

$$f(x) = e^{-x^2}$$

または

$$f(x) = e^{-ax^2} \quad (a > 0)$$

という分布に近づいていきます.これをガウス関数といいます.この関数のグラフを図 16 に示します.

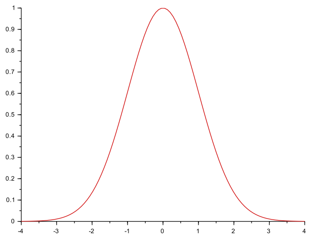


Fig.16: e^{-x^2}

ガウス関数の特性を見ていくと,まず, $x = 0$ を代入すると

$$f(0) = e^0 = 1$$

となります.次に

$$f(-x) = e^{-(-x)^2} = e^{-x^2} = f(x)$$

の性質を持っているので偶関数です.よって, y 軸に関して左右対称になります.ガウス関数を微分すると

$$f'(x) = -2xe^{-x^2}$$

$x < 0$ では単調増加, $x > 0$ では単調減少なので, x の絶対値が増えていくと, $f(x)$ の値は小さくなっていきます.また,極限

$$\lim_{x \to \pm\infty} e^{-x^2} = 0$$

よって,中心から無限遠で 0 になります.

次に e^{-ax^2} の a を変化させたときのグラフを図 17 に示します.

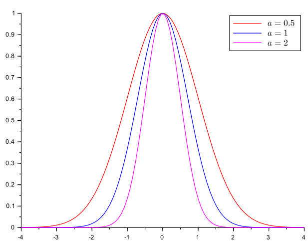


Fig.17: e^{-ax^2}

a の値を大きくしていくと,より中心に集中していくことがわかります.また, a の値を小さくすれば,分布が拡がっていくこともわかります.つまり, a は分布の大きさを表しています.

ここで, ガウス関数を確率密度関数に変換します.まず,

$$\int_{-\infty}^{\infty} e^{-ax^2} dx = 1$$

を満たす必要があります.ガウス積分から

$$\int_{-\infty}^{\infty} e^{-ax^2} dx = \sqrt{\frac{\pi}{a}}$$

となるので

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

と正規化します.また, この a は分布の大きさですが, 分布のばらつきを表すものに分散があります.そこで

$$2\sigma^2 = \frac{1}{a} \quad \therefore a = \frac{1}{2\sigma^2}$$

と置き換えます.すると

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

となります.この関数は $x = 0$ を中心とした分布ですが, 図 15 でわかるように中心は移動します.そこで中心がどの位置になるかというと平均 μ になります.この対応をすると

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

この確率密度関数の分布をガウス分布または正規分布といいます.この分布は平均 μ と分散 σ^2 によって形が決まるので, 正規分布を $N(\mu, \sigma^2)$ と表します.

平均 $\mu = 0$ としたとき, 分散 σ^2 の値によって正規分布のグラフがどう変化するかを図 18 に示します.

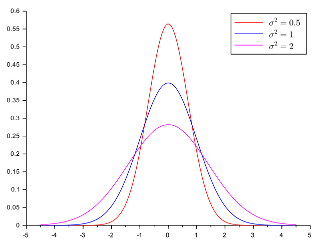


Fig.18: $N(0, \sigma^2)$

正規分布に従った確率変数 X があるとき, ある範囲 $(a \leq X \leq b)$ の確率は

$$P(a \leq X \leq b) = \int_a^b \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

で求めることができます.この計算をするために, この式を変形していきます.まず, 変数変換 $t = x - \mu, \quad dt = dx$ をすると

$$\int_{a-\mu}^{b-\mu} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{t^2}{2\sigma^2}} dt$$

として, 中心の位置を 0 にします.さらに, 次の変数変換をすると

$$z = \frac{t}{\sigma} \qquad dz = \frac{dt}{\sigma}$$

積分範囲を

$$\alpha = \frac{a-\mu}{\sigma}, \qquad \beta = \frac{b-\mu}{\sigma}$$

とおくと

$$\int_{\alpha}^{\beta} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$$

これは $\sigma = 1$ にしています.この 2 回の変換をまとめると

$$z = \frac{x-\mu}{\sigma}$$

とすることができます.この変数に対応した確率密度関数は

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

となります.これは平均 $\mu = 0$, 分散 $\sigma^2 = 1$ に相当します.つまり, この分布は $N(0,1)$ であり, このような正規分布を標準正規分布といいます.また, この変換のことを標準化変換, z を標準化変数と呼ぶこともあります.

すべての正規分布は

$$z = \frac{x-\mu}{\sigma}$$

という変数変換をすると標準正規分布となり,

$$x = \sigma z + \mu$$

という逆変数変換をすると, 元の正規分布に戻すことができます.積分を計算するときは, 標準正規分布に変換して, 標準正規分布表を使う方法や, 指数関数のべき級数展開

$$\int_0^a e^{-ax^2} dx = a - \frac{1}{3\cdot 1}a^3 + \frac{1}{5\cdot 2!}a^5 - \frac{1}{7\cdot 3!}a^7 + \frac{1}{9\cdot 4!}a^9 - \cdots$$

を利用する方法があります.また, 正確な値が必要でなければ累積分布関数の近似式を使うこともできます.

$$F(x) = \frac{1}{1+\exp(-0.07056x^3 - 1.5976x)}$$

この関数の誤差は 0.0014 以下で

$$F(x) = \frac{1}{1 + \exp(-1.7x)}$$

こちらは誤差が 0.01 以下になります. $a \leq x \leq b$ の範囲を求める場合は $F(b) - F(a)$ で求められます. 1 つ目の近似式は一般の正規分布でも同じ誤差になります.

$$F(x) = \frac{1}{1 + \exp\left(-0.07056\left(\frac{x-\mu}{\sigma}\right)^3 - 1.5976\left(\frac{x-\mu}{\sigma}\right)\right)}$$

正規分布 $N(\mu, \sigma^2)$ の性質として, 一般に確率変数は平均 μ から $\pm\sigma$ の間に 68%, $\pm 2\sigma$ の間に 95%, $\pm 3\sigma$ の間に 99.7% 以上存在します.

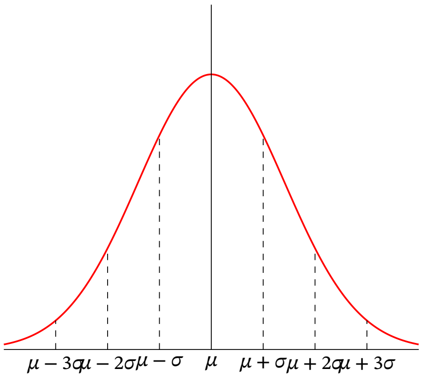


Fig.19: 正規分布の性質

例えば, $\mu \pm 3\sigma$ 以外のところ(0.03%)は誤差として扱うというように, 区切りとして $\mu \pm \sigma$, $\mu \pm 2\sigma$, $\mu \pm 3\sigma$ がよく使われます.

ここで, はじめに言ったことを振り返って考えてみると, 2 項分布の n を大きくして標準化すれば標準正規分布 $N(0, 1)$ に近づいていくということでした. 実は, 2 項分布だけでなく, 平均が μ, 分散が σ^2 である他の分布に従っている確率変数 $X_1, X_2, \ldots, X_n$ の平均を $\overline{X}$ として

$$Z_n = \frac{\sqrt{n}}{\sigma}(\overline{X} - \mu)$$

とすると, n を大きくしたとき, Z_n の分布は標準正規分布に近づいていくことがわかっています(ただし, すべての分布で成立するとは限りません). これを中心極限定理といいます.

22. モーメント

ここでは詳しい説明をせずに軽く見ていくことにします.

$f(x)$ を確率密度関数として, $\phi(X)$ を確率変数 X の関数とします. $\phi(X)$ の期待値 $E[\phi(X)]$ は

$$E[\phi(X)] = \int_{-\infty}^{\infty} \phi(x) f(x) dx$$

で与えられます. 離散でも同じように考えることができます. このとき,

$$E[X] = \int_{-\infty}^{\infty} x f(x) dx, \qquad E[X^2] = \int_{-\infty}^{\infty} x^2 f(x) dx, \qquad E[X^3] = \int_{-\infty}^{\infty} x^3 f(x) dx$$

となり, $\phi(X) = X^k (k = 0, 1, 2, \ldots)$ とすると, 一般式は

$$E[X^k] = \int_{-\infty}^{\infty} x^k f(x)dx$$

で与えられます.これを k 次のモーメントといいます.よって，1 次のモーメント

$$E[X] = \int_{-\infty}^{\infty} xf(x)dx = \mu$$

は平均であり，

$$E[(X-\mu)^2] = \int_{-\infty}^{\infty} (X-\mu)^2 f(x)dx = \sigma^2$$

は平均のまわりの 2 次モーメントといいます.また $\mu = 0$ とすれば，分散になります.そして

$$E[(X-\mu)^3] = \int_{-\infty}^{\infty} (X-\mu)^3 f(x)dx = \gamma$$

は平均のまわりの 3 次モーメントで，歪度といい，分布のゆがみの大きさです.さらに 4 次モーメントは尖度といい，中心の周囲の部分の尖り具合を表します.

ここで，

$$\phi(X) = e^{tx}$$

としたとき，

$$M(X) = E\left[e^{tx}\right]$$

をモーメント母関数といいます.指数関数の級数展開

$$e^x = 1 + x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + \cdots + \frac{1}{n!}x^n + \cdots$$

から，e^{tx} を考えると

$$e^{tx} = 1 + tx + \frac{1}{2!}t^2x^2 + \frac{1}{3!}t^3x^3 + \cdots + \frac{1}{n!}t^nx^n + \cdots$$

と展開できます.この関数の期待値は

$$E\left[e^{tx}\right] = 1 + E[x]t + \frac{1}{2!}E[x^2]t^2 + \frac{1}{3!}E[x^3]t^3 + \cdots + \frac{1}{n!}E[x^n]t^n + \cdots$$

となります.これを t で微分すると

$$\left(E\left[e^{tx}\right]\right)' = E[x] + \frac{1}{2!}E[x^2]t + \frac{1}{3!}E[x^3]t^2 + \cdots + \frac{1}{(n-1)!}E[x^n]t^{n-1} + \cdots$$

これに $t = 0$ を代入すると $E[x]$ が求まります.同様に，t でまた微分して $t = 0$ を代入すると $E[x^2]$ が得られます.つまり，モーメント母関数を微分すれば，k 次のモーメントを求めることができるので，モーメント母関数という名前になっています.一般に k 次モーメントは

$$E[x^k] = M^{(k)}(0)$$

で与えられます.

23. 最後に

いかがだったでしょうか．なるべくわかりやすく書いたつもりですが，わかりづらかったところがあったかもしれません．また，知識不足で間違っていたり，説明足らずなところもあったかもしれません．気づいたところがあれば気軽にご連絡してもらえると助かります．内容についてはまだまだ足りておらず，取り上げなかった分布も多く，相関やランダムウォーク，マルコフ過程とかに全然到達できませんでした．また，CG ということでレンダリングで使われている確率分布(Phong 分布，Beckmann 分布，GGX 分布，GTR 分布，AGC 分布，SGD 分布など)にも触れられませんでした．まあ，これらは次の機会ということで．少しでも誰かの参考になれば幸いです．

参考文献

[1] 村上雅人「なるほど確率論」海鳴社, 2003

[2] 村上哲哉「確率」ファーラム・ A, 1989

[3] 薩摩順吉「確率・統計」岩波書店, 1989

[4] 東京大学教養部統計学教室「統計学入門」東京大学出版会, 1991

[5] 竹内淳「高校数学でわかる統計学」講談社, 2012