

MAT220 第8回講義ハンドアウト

二項分布 第I部

目次

1 復習：数え上げと分布	2
1.1 3回のコイン投げで2回だけ裏が出る確率を計算する	2
1.2 この問題の一般的な解法	3
1.3 確率分布	4
1.4 大数の法則	6
2 例：二項実験	7
2.1 例1：公平なコインを10回投げる	7
2.2 他の n の値に対する二項分布	9
2.3 例2：不公平なコインを投げる	10
3 一般論	11
3.1 二項実験の定義	11
3.2 p は分布に何をするのか	12
3.3 二項分布の期待値と分散	13
4 練習問題	14
4.1 ルーレット盤	14
4.2 ルートボックス	16
4.3 回転寿司	16
4.4 サイコロを振る	17
4.5 バスケットボールをする猫	18

この講義では、確率論についての議論の最後の方のトピックの一つである二項実験を扱います。これから見るように、二項実験とは、起こり得る結果が二つだけ、すなわち「成功」または「失敗」だけであるような特別な設定です。可能性が二つしかないので、このようなシステムはかなり記述しやすく、計算もしやすいです。そのため、これまでに学んだ内容と簡単に結びつけることができます。一般論を説明した後、二項実験のいくつかの例を見ていきます。

1 復習：数え上げと分布

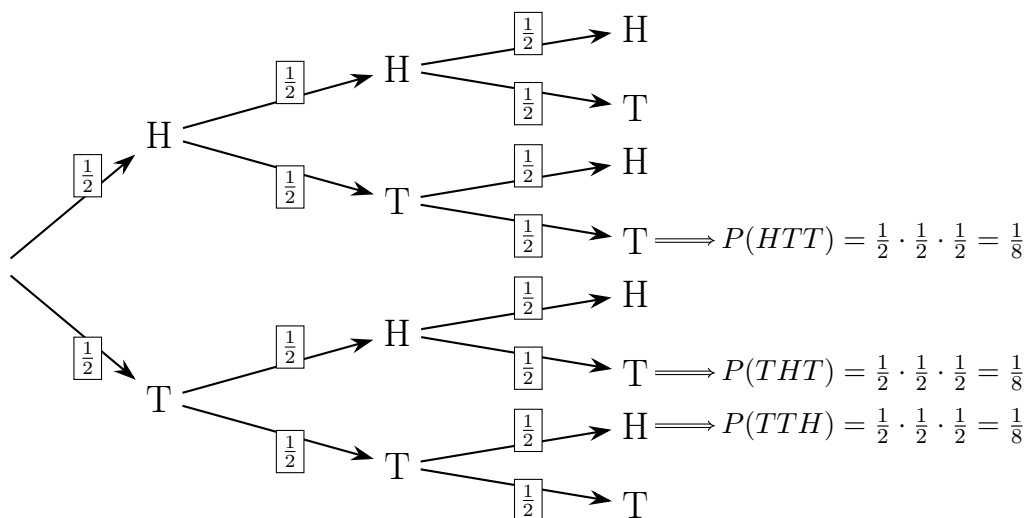
1.1 3回のコイン投げで2回だけ裏が出る確率を計算する

公平なコインがあり、それを3回投げるとします。ちょうど2回だけ裏が出る確率はいくらでしょうか。幸い、3のような小さい数であれば、可能な結果をすべて書き出すことができます。

- 1回のコイン投げには2つの結果があります： H または T 。
- 2回のコイン投げには4つの結果があります： HH, HT, TH, TT 。
- 3回のコイン投げには8つの結果があります（完全な標本空間）：

$$\{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}.$$

P (ちょうど2個の T)のような確率を計算する方法はいくつかあります。一つの方法は、（第7回講義で見たような）決定木を使って、それぞれの有利な結果の確率を計算することです。



HTT, TTH, THT という三つの事象は互いに排反なので、これらの個別の結果の確率を足し合わせれば、複合事象の確率が得られます。

$$P(\text{ちょうど2個の } T) = P(HTT \text{ または } THT \text{ または } TTH) = \frac{1}{8} + \frac{1}{8} + \frac{1}{8} = \frac{3}{8}.$$

あるいは、（第5回講義で見たような）複合事象のより簡単な計算を使って、同じ答えにたどり着くこともできます。この状況ではすべての結果が同様に起こりやすいので、黄金公式を直接使用することもできます。

$$P(\text{事象}) = \frac{\text{その事象に有利な結果の数}}{\text{結果の総数}}.$$

この場合、ちょうど2回裏を含む結果は HTT, THT, TTH の3つなので、

$$P(\text{ちょうど2個の } T) = \frac{\text{2個の } T \text{ を含む結果の数}}{\text{結果の総数}} = \frac{3}{8}.$$

解答

結果の総数は $2^5 = 32$ です。ちょうど2回裏が出る結果の数は $C_{5,2} = 10$ です（パスカルの三角形の第5行）。したがって、

$$P(\text{ちょうど2回裏}) = \frac{C_{5,2}}{2^5} = \frac{10}{32} = \frac{5}{16}.$$

1.3 確率分布

変数 X が確率変数と呼ばれるのは、ある実験または観測において、 X が取る値がランダムな結果によって決まるときです。主な種類は二つあります。

- 離散型確率変数は、有限個の値、または可算個の値だけを取ることができます。
- 連続型確率変数は、区間内の無数の値のどれでも取ることができます。

練習：離散型か連続型か

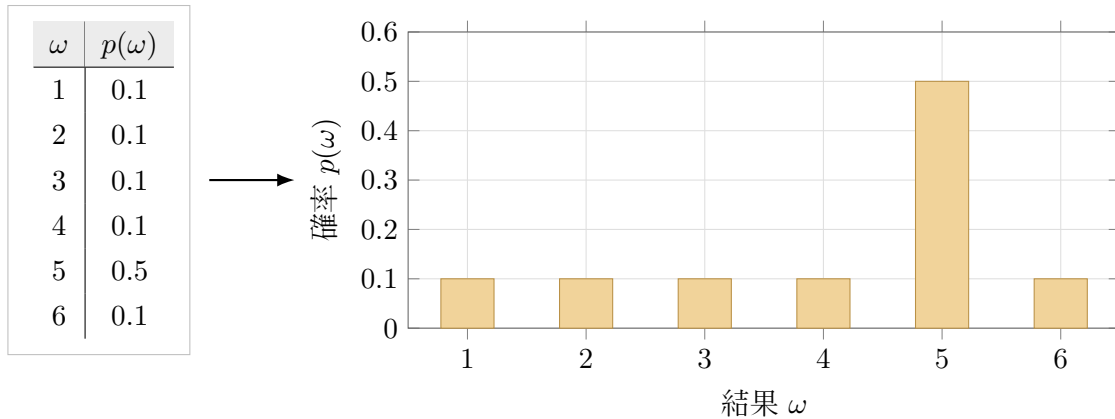
それぞれの確率変数を、離散型または連続型に分類しなさい。

- (a) 選ばれた学生が秋学期の登録を完了するまでにかかる時間。
- (b) ランダムに選ばれた日に、選ばれた学生が受信するテキストメッセージの数。
- (c) 電気自動車が満充電で走行できるマイル数。
- (d) ある地区の登録有権者から50人をランダムに選び、前回の郡選挙で投票した人数を求める。

解答

- (a) 連続型。
- (b) 離散型。
- (c) 連続型。
- (d) 離散型。

離散型確率分布とは、離散型確率変数のそれぞれの異なる値に確率を割り当てたものです。第6回講義では、確率分布を棒グラフとして可視化できることを見ました。例えば、次のグラフは、



確率が可能な値1, 2, 3, 4, 5, 6の間にどのように広がっているかを視覚的に説明しています。

通常、確率変数は X や Y のように大文字で書きます。もし X が1, 2, 3, 4, 5, 6という値を取ることができるなら、

$$X = 4 \quad \text{または} \quad X = 5$$

のような文は、確率変数の可能な値を表しています。 X が値4を取る確率は、次のように書けます。

$$P(X = 4).$$

確率分布では、これは4の位置にある棒の高さで表されます。上の例では、

$$P(X = 4) = 0.1.$$

また、 X がある範囲の値に入る確率を尋ねることもできます。例えば、

$$P(X < 3)$$

は、 X が3より小さい値を取る確率を意味します。3より小さい値は1と2なので、合計確率は和を取ることで計算できます。

$$P(X < 3) = P(X = 1) + P(X = 2) = 0.1 + 0.1 = 0.2.$$

第6回講義では、確率変数の期待値の公式も見ました。これは確率分布の平均のように働きます。確率的システムの期待値は加重平均です。有限個の値を取る離散型確率変数について、期待値は

$$E(X) = \sum_{i=1}^n p_i x_i = p_1 x_1 + \cdots + p_n x_n.$$

です。この公式は、確率分布の幾何学的な重心を計算しようとすることで慎重に導くことができます。詳しい説明は第6回講義を参照してください。非公式には、 $E(X)$ は X の長期的に期待される振る舞いです。

確率分布の分散は、 X の可能な値が期待値のまわりにどれだけ広がっているかを測るものです。これは次のように定義されます。

$$\text{Var}(X) = E((X - E(X))^2).$$

この公式が、データセットに対する母分散の通常の変換と似ていることに注目してください。まとめると、次のようになります。

期待値と分散

離散型確率変数の期待値は

$$E(X) = \sum_{i=1}^n p_i x_i.$$

です。これは平均の確率論版であり、したがって分布の中心を測るものです。

確率変数の分散は

$$\text{Var}(X) = E((X - E(X))^2).$$

です。これは、分布が期待値のまわりにどれだけ広がっているかを測ります。

1.4 大数の法則

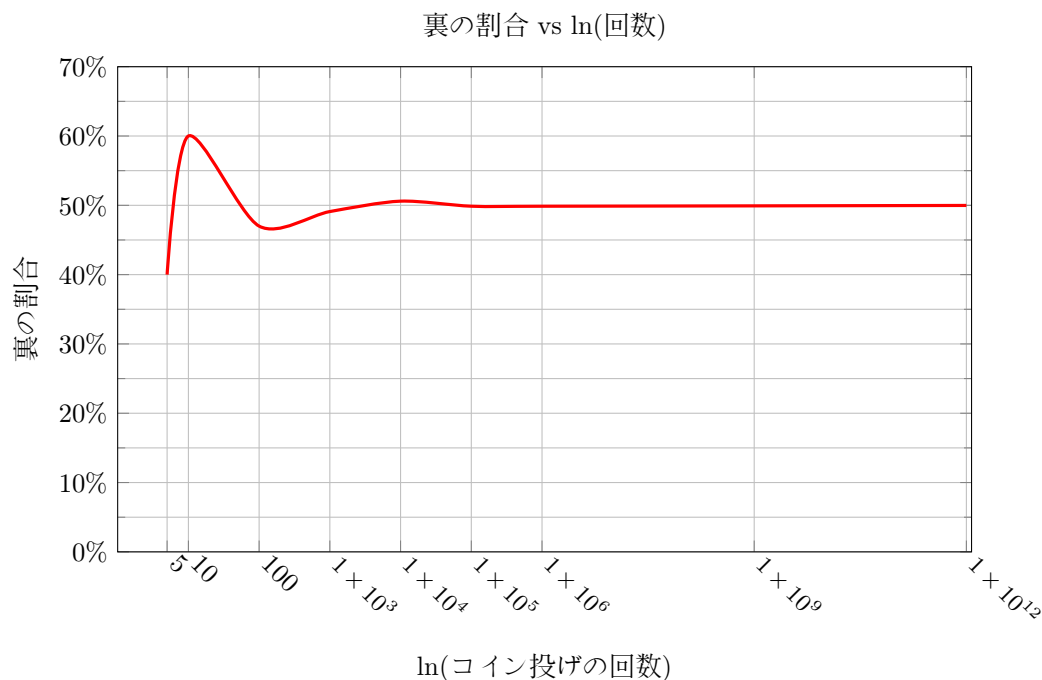
確率変数の期待値は、実験の平均的で長期的に期待される結果を計算する方法です。ここでは、確率の別の解釈を導入することで、この考えをさらに詳しく説明します。

大数の法則は、ある実験を何度も繰り返すと、平均的な結果が理論上の平均に近づいていく、というものです。例として、公平なコインを考えましょう。コインを投げて裏が出る確率は50%、つまり $P(T) = \frac{1}{2}$ であると期待します。私たちの確率の解釈では、長期的には裏が約50%出ると期待することになります。

これを見るために、コインを何度も投げ始めたとします。乱数生成器を用いた例として、コイン投げの標本における裏の割合を以下に示します。

- 5回：HHTTH。 $T = 2$, $H = 3$ なので、裏は40%。
- 10回：HHTTHTHTTT。 $T = 6$, $H = 4$ なので、裏は60%。
- 100回： $T = 47$, $H = 53$ なので、裏は47%。
- 1000回： $T = 491$, $H = 509$ なので、裏は49.1%。
- 10000回： $T = 5062$, $H = 4938$ なので、裏は50.6%。
- 100000回： $T = 49872$, $H = 50128$ なので、裏は49.87%。
- 1百万回：裏は49.87%。
- 1十億回：裏は49.92%。
- 1兆回：裏は49.982%。

これをグラフにすると次のようになります。ここでは、コイン投げの回数が同じ軸に収まるように、対数スケール $\ln(\cdot)$ を使っています。



見て分かるように、最初のうちは結果にかなり大きなランダムな変動があります。しかし、時間がたつにつれて、これらのランダムな変化はほとんど打ち消し合い、裏の割合は理論上の平均である50%に収束していきます。これは、確率の一つの解釈を与えてくれます。

確率の頻度主義的解釈

$P(A) = x\%$ と言うとき、もしその実験を非常に多くの回数繰り返したなら、事象Aに対応する結果の長期的な割合は $x\%$ にだんだん近づいていくと期待する、という意味です。

この意味で、期待値 $E(X)$ は、独立な条件のもとで実験を何度も繰り返したときに得られる平均的な結果として、よりよく理解できます。最終的には、この繰り返しの活動の移動平均は何かに収束していき、その何かが期待値です。後で見るように、期待値は二項実験について議論するときに特に役立ちます。

2 例：二項実験

2.1 例1：公平なコインを10回投げる

公平なコインがあり、それを続けて10回投げるとします。ちょうど5回裏が出る確率はいくらでしょうか。

解答

前の節に基づいて、次の公式を使って $P(\text{ちょうど5回裏})$ を計算できます。

$$P(\text{ちょうど5回裏}) = \frac{C_{10,5}}{2^{10}}.$$

もちろん、通常の公式 $C_{n,r} = \frac{n!}{r!(n-r)!}$ を使ってもよいですが、説明の流れをそろえるために、ここではパスカルの三角形を使います。三角形の第10行は次の通りです。

1 10 45 120 210 252 210 120 45 10 1

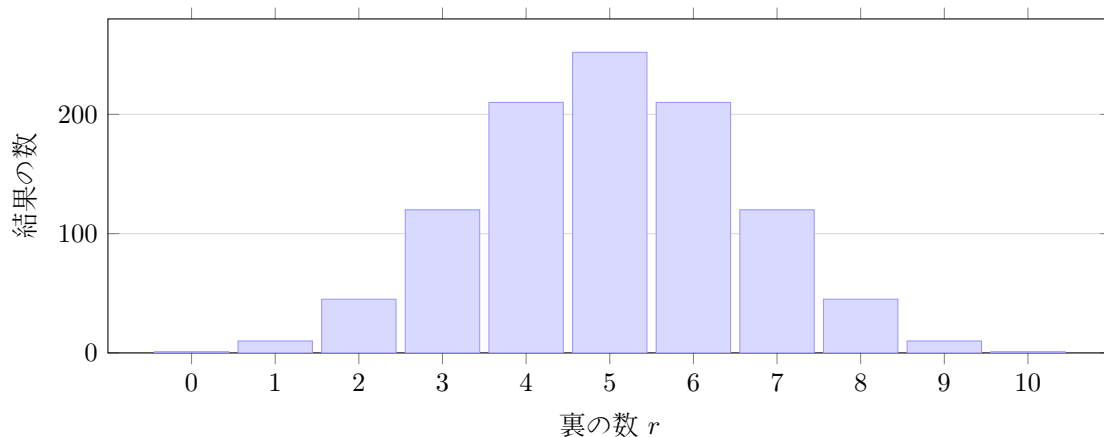
したがって、 $r=5$ に対応する項まで進むと（三角形は0から数え始めることを思い出してください）、 $C_{10,5} = 252$ であることが分かります。結果の総数は $2^{10} = 1024$ です。したがって、

$$P(5 \text{ 回裏}) = \frac{C_{10,5}}{2^{10}} = \frac{252}{1024} \approx 0.246 \approx 24.6\%.$$

ここで、10回のコイン投げの標本空間の構造を調べてみましょう。上の議論から、この標本空間には1024個の要素があることが分かっています。これらは $HHHHHHHHHH$ から $TTTTTTTTTT$ まで、そしてその間のすべての結果を含みます。この大きな空間を整理する自然な方法は、結果に含まれる裏の数によってグループ分けすることです。¹ 標本空間を理解する便利な方法は、結果に含まれる裏の数でグループ分けすることです。パスカルの三角形の第10行に基づく、1024個の結果からなる標本空間は次のように整理できます。

裏の数 r	0	1	2	3	4	5	6	7	8	9	10
結果の数	1	10	45	120	210	252	210	120	45	10	1

上の表は棒グラフとして可視化できます。



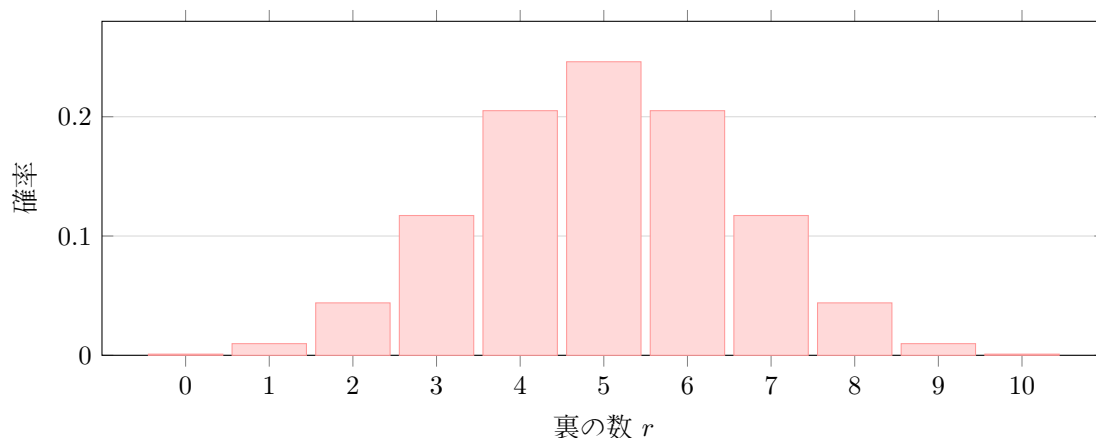
ここで、 x 軸は離散的な個数、つまり裏の数を表しています。 y 軸は対応する結果の数です。したがって、ある意味で、このグラフはパスカルの三角形の第10行を可視化したものです。

¹もちろん、表の数から始めても同じことができます。しかし、ここまでずっと裏を使ってきたので、ここでも裏を使います。

上の議論に基づいて、このシステムにおける確率は $C_{10,r}$ を $2^{10} = 1024$ で割ることで計算できることが分かっています。したがって、その割り算を行い、その結果をプロットすることで、このシステムで確率そのものがどのように分布しているかを見ることができます。

r 回裏	0	1	2	3	4	5	6	7	8	9	10
確率 (分数)	$\frac{1}{1024}$	$\frac{10}{1024}$	$\frac{45}{1024}$	$\frac{120}{1024}$	$\frac{210}{1024}$	$\frac{252}{1024}$	$\frac{210}{1024}$	$\frac{120}{1024}$	$\frac{45}{1024}$	$\frac{10}{1024}$	$\frac{1}{1024}$
近似値 (割合)	0.1%	1.0%	4.4%	11.7%	20.5%	24.6%	20.5%	11.7%	4.4%	1.0%	0.1%

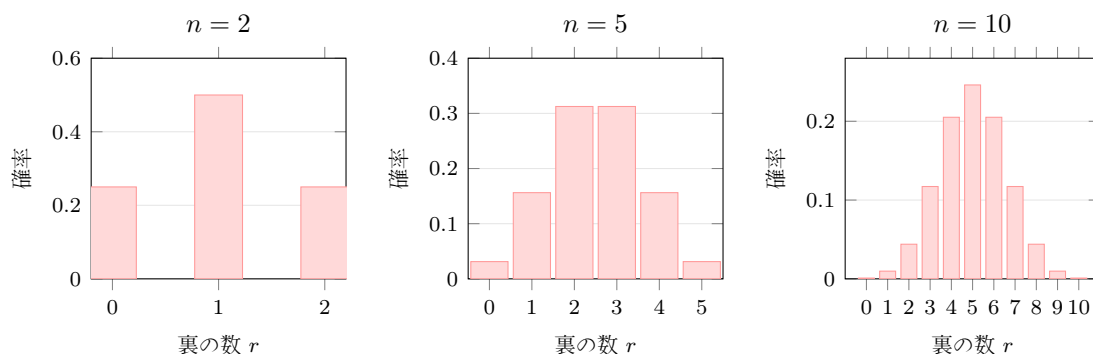
先ほど得た24.6%という値に注目してください。上の表を使って、この確率分布をグラフにできます。

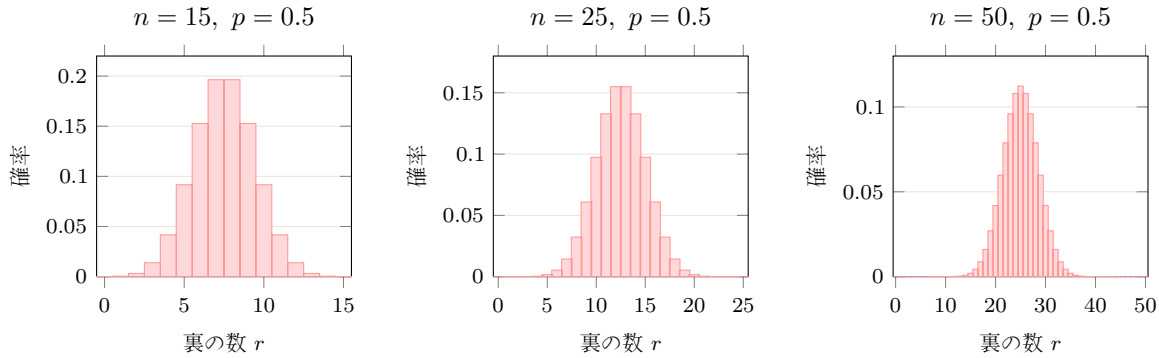


この形は、先ほどの棒グラフの形と完全に同じであることに注意してください。変わったのはy軸のスケールだけです。この特別なグラフには名前があります。二項分布と呼ばれます。

2.2 他の n の値に対する二項分布

一般に、コインを投げる総回数 n の別の値に対しても、二項分布を作ることができます。これらのグラフの形は、コイン投げの非常に多くの可能な結果に確率がどのように分布しているかを表します。その形はパスカルの三角形の第 n 行から生じ、 2^n で割ると、それらの数が確率になります。一般に、コイン投げを続けて n が大きくなると、 r の可能な値が増え、分布はより滑らかになり、中央付近 ($r = n/2$ の近く) により集中していきます。この「滑らかさ」については、第10回講義で詳しく話します。





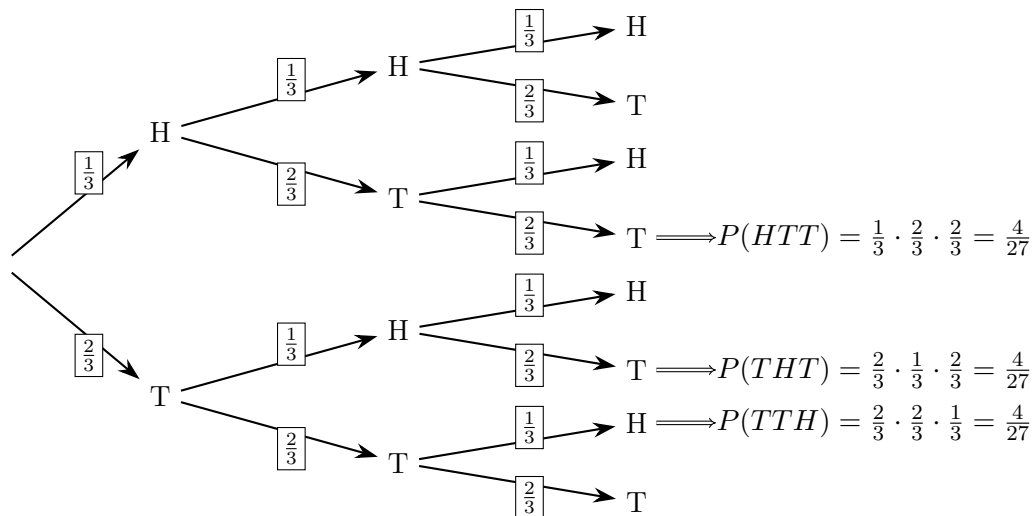
2.3 例2：不公平なコインを投げる

今度は、何らかの仕方で細工されたコインがあり、表または裏が出る確率がもはや等しくないとしてます。この「不公平な」コインの確率を次のように仮定します。

$$P(H) = \frac{1}{3}, \quad P(T) = \frac{2}{3}.$$

問題：3回投げたとき、ちょうど2回裏が出る確率はいくらでしょうか。

上で議論した公平なコインの場合には、8つの結果がすべて同様に起こりやすかったので、有利な結果の数（3つ）を数えて、それを8で割るだけでよかったです。しかし、不公平なコインでは、すべての結果が同様に起こりやすいわけではありません。したがって、有利な結果を数えて8で割るだけではいけません。代わりに、決定木の方法を使うことができます。可能な結果の空間は前と同じですが、木の各枝に付く確率が変わっています。



ちょうど2回裏が出ることは、 HTT , THT , TTH という結果に対応します。これらは互いに排反な事象なので、求めたい確率は和を取ることで得られます。

$$P(\text{ちょうど2回裏}) = P(HTT \text{ または } THT \text{ または } TTH) = \frac{4}{27} + \frac{4}{27} + \frac{4}{27} = \frac{12}{27}.$$

それぞれの結果の確率計算は本質的に同じであることに注意してください。 $P(HTT)$, $P(THT)$, $P(TTH)$ を計算するには、 $P(T) = \frac{2}{3}$ を二つ掛け、 $P(H) = \frac{1}{3}$ を一つ掛ければよいです。実際、こ

の計算の仕方はさらに一般の確率を計算する場合にも使えます。 n 回投げるとき、次のようになります。

$$P(r \text{ 個の } T) = C_{n,r} \left(\frac{2}{3}\right)^r \left(\frac{1}{3}\right)^{n-r}.$$

さらに一般に、裏または表が出る確率を別の値に変えたとしても、 n 回のコイン投げで r 回裏が出る確率は次の公式で計算できます。

$$P(n \text{ 回中 } r \text{ 個の } T) = C_{n,r} (P(T))^r (P(H))^{n-r}.$$

ここで、 $(P(T))^r (P(H))^{n-r}$ は、 r 個の T と $(n-r)$ 個の H を含む特定の一つの順序の確率であり、 $C_{n,r}$ は r 個の T と $(n-r)$ 個の H の異なる順序の数を数えています。

3 一般論

3.1 二項実験の定義

第2節では、コイン投げの二つの異なる場合を見ました。一つは公平な場合で、もう一つは二つの可能な結果が異なる確率を持つという意味で「不公平」な場合です。実は、これらは二項実験と呼ばれるより一般的な過程の二つの例です。二項実験には四つの重要な特徴があります。

定義：二項実験

二項実験は次の性質を持ちます。

- (B1) 試行回数が固定されている。この数を n と呼ぶ。
- (B2) n 回の試行は独立であり、同じ条件のもとで繰り返される。
- (B3) 各試行には二つの結果しかない：成功 (S) と失敗 (F)。
- (B4) 各試行で成功する確率は同じである。これを p と呼び、失敗する確率を q と呼ぶ。各試行は成功または失敗のどちらかなので、 $p + q = 1$ であり、したがって $q = 1 - p$ である。

言い換えると、二項実験とは、独立な試行を固定された回数だけ行い、各試行には二つの可能な結果があり、成功確率が同じであるような実験です。二項実験における中心的な問題は、 n 回の試行のうち r 回成功する確率を求めることです。仮定(B2)は、これらの連続した確率の計算を簡単にするために使われます。

3.1.1 例：公平なコイン投げ

前の節の最初の例では、次のようになっていました。

- (B1) n はコインを投げる回数でした。
- (B2) コイン投げは独立です。実際、それらは前の実験の結果とはまったく関係していないものとして扱います（物理的にも）。
- (B3) 成功は「裏が出た」と定義され、失敗は「裏が出なかった」と定義されました。
- (B4) 公平なコインでは、 $p = P(T) = 0.5$ かつ $q = 0.5$ です。

この実験に関する確率を求めるという中心的な問題については、まさに前の議論で扱っていたことです。すでに、公平なコインについて

$$P(n\text{回中}r\text{回裏}) = \frac{C_{n,r}}{2^n}$$

であることを見ました。

不公平なコインについては、二項実験としての設定は本質的に同じですが、確率が $p = \frac{2}{3}$ と $q = \frac{1}{3}$ に変わっています。先ほど、異なる r の値に対応する確率は次の公式で計算できることを見ました。

$$P(r \text{ 個の } T) = C_{n,r} \left(\frac{2}{3}\right)^r \left(\frac{1}{3}\right)^{n-r} = C_{n,r} \cdot P(T)^r \cdot P(H)^{n-r}.$$

二項確率の一般公式

パラメータ $n, p, q = 1 - p$ を持つ二項実験において、ちょうど r 回成功する確率は

$$P(n\text{回の試行中}r\text{回成功}) = C_{n,r} p^r q^{n-r}.$$

です。ここで、

- p^r は、特定の一つの順序における確率の成功部分です。
- q^{n-r} は、その同じ特定の順序における確率の失敗部分です。
- $p^r q^{n-r}$ は、 r 回の成功と $(n-r)$ 回の失敗を含む特定の一つの列の確率です。
- $C_{n,r}$ は、 r 回の成功と $(n-r)$ 回の失敗を含む異なる列がいくつあるかを数えます。

3.2 p は分布に何をするのか

一般に、公式

$$P(n\text{回の試行中}r\text{回成功}) = C_{n,r} p^r q^{n-r}.$$

を使って、すべての二項実験に対応する二項分布を記述できます。簡単に言えば、 x 軸を確率変数の可能な値（0から n まで）とし、 y 軸を上の公式の出力とします。

公平なコインの場合、成功確率は $p = 0.5$ で、これは q とちょうど等しいです。このため、得られる二項分布は完全に対称になります。第2節のグラフでこれに注目してください。成功確率を $p = 0.5$ からずらし始めると、分布は対称性を失います。

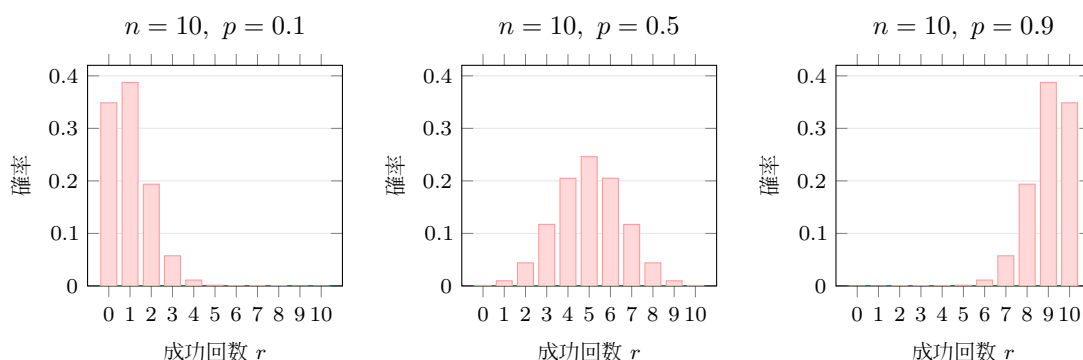
説明のために、 $n = 10$ 回の試行からなる二項実験を考えているとしましょう。確率が $p = 0.5$ の場合、得られる分布は第2.1節の図です。ここでは、 p を大きくしたり小さくしたりすると何が起こるかを簡単に説明します。

- p が0に近い場合、例えば $p = 0.1$ の場合、各試行で成功する確率はかなり低くなります。したがって、10回の試行で7, 8, 9, 10回のような多くの成功を得る確率は非常に低いです。最も起こりやすいのは、10回の試行で0, 1, または2回程度の成功です。そのため、このようなシステ

μに対応する二項分布では、確率の大部分が x 軸の左側、つまり0, 1, 2に近い値の周辺に分布します。

- p が1に近い場合、例えば $p = 0.9$ の場合、今度は各試行で成功することが非常に起こりやすくなります。したがって、10回の試行で成功が1回または2回しかないのは非常に運が悪い場合であり、むしろ7, 8, 9, 10回程度の成功を期待することになります。二項分布のレベルでは、このように成功確率が高いと、分布は右側に押し出され、確率の大部分は x 軸上の8, 9, 10付近に集中します。

下のグラフは、 $n = 10$ 回の試行に対して、成功確率がそれぞれ0.1, 0.5, 0.9である場合の実際の二項分布を表しています。



3.3 二項分布の期待値と分散

二項分布には、特に重要な公式が二つあります。

二項分布の期待値と分散

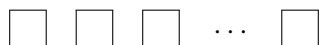
X が、パラメータ $n, p, q = 1 - p$ を持つ二項実験における成功回数を数えるなら、

$$E(X) = np, \quad \text{Var}(X) = npq.$$

です。

これらの公式が便利なのは、分布全体を手で計算しなくても、二項分布の中心と広がりをすばやく理解できるからです。

技術的には、上の二つの公式は代数を使って確認できます。しかし、ここではそれを行わず、大まかに議論します。 n 回の試行を n 個の枠として想像してください。



各枠は、成功1回または成功0回を寄与します。各枠で成功する確率は p なので、確率の頻度主義的解釈によれば、各枠は長期的な期待平均として p 回の成功を寄与します。枠は n 個あるので、成功回数の期待値は

$$p + p + \cdots + p = np,$$

となり、公式 $E(X) = np$ が得られます。分散の公式については、ここでは導きません。しかし、公式

$$\text{Var}(X) = npq$$

は成功回数がどれだけ変動しやすいかを教えてください。

- p が0に近い、または1に近い場合、結果がより予測しやすいので、実験はあまり変動せず、 $E(X)$ のまわりにより密に集まります。
- p が0.5に近い場合、成功と失敗のどちらも各試行でかなり起こりやすいので、実験はより変動しやすくなります。

4 練習問題

4.1 ルーレット盤

私たちのクラスに20人の学生がいるとします。あなたは一生懸命勉強し、数年後に無事LUJを卒業します。Dr. O'Connellはあなたたちをととても誇りに思い、全員をカジノに連れて行き、ルーレット盤で遊ばせることにしました。各学生はルーレット盤を1回ずつ回します。



ルーレット盤には36個の番号付きの枠があり、緑の枠を含めると合計で37個の結果があります。Dr. O'Connellはあなたに次の提案をします。

ボールが緑の枠に止まったら\$50,000をもらい、それ以外の結果なら\$0をもらう。

答えたい主な問題は次の通りです。

練習

ちょうど3人の学生が勝つ確率はいくらでしょうか。

この問題を解くために、第3.2節で述べた一般的な二項公式を使います。しかし、何かを計算し始める前に、まずこの状況が本当に二項実験であることを確認しなければなりません。

- B1: 20人の学生がいて、それぞれがルーレット盤を1回ずつ試すので、 $n = 20$ 回の試行があります。
- B2: 各回転が独立であり、ルーレット盤が公平であると仮定すると、試行は確かに互いに独立です。
- B3: 各試行には二つの可能な結果しかありません。

成功 = 緑に止まる, 失敗 = それ以外に止まる.

- B4: ルーレット盤の各枠は、結果になる可能性が等しいです。盤には37個の枠があり、緑の枠は一つだけなので、成功と失敗の確率は

$$p = P(\text{成功}) = \frac{1}{37}, \quad q = P(\text{失敗}) = 1 - p = \frac{36}{37}.$$

です。

これで、このシステムを二項実験として見られることを確認できたので、主な問題を解くことができます。

解答

二項公式を使います。

$$P(n\text{回の試行中}r\text{回成功}) = C_{n,r} p^r q^{n-r}.$$

この場合、学生が20人いるので $n = 20$ 、 $r = 3$ 、 $p = \frac{1}{37}$ 、 $q = \frac{36}{37}$ です。これらの値を公式に代入すると、

$$P(\text{ちょうど3人が勝つ}) = C_{20,3} \left(\frac{1}{37}\right)^3 \left(\frac{36}{37}\right)^{17} = \frac{20!}{3! \cdot 17!} \left(\frac{1}{37}\right)^3 \left(\frac{36}{37}\right)^{17}.$$

数値的には、これはおよそ

$$P(\text{ちょうど3人が勝つ}) \approx 0.014 \approx 1.4\%.$$

です。

この確率の別の説明として、私たちは20回の独立な試行を行い、3回の成功と17回の失敗を求めています。これを表す一つの特定の結果は、例えば

$$S, S, S, F, F, \dots, F,$$

です。つまり、最初の3人の学生が成功し、その後の17人の学生がすべて失敗する場合です。仮定(B2)により、これらの試行はすべて互いに独立です。したがって、この結果に対応する確率は、それぞれの結果の確率を単に掛け合わせることで計算できます。つまり、

$$P(S \text{ then } S \text{ then } S \text{ then } F \text{ then } \dots \text{ then } F) = \frac{1}{37} \cdot \frac{1}{37} \cdot \frac{1}{37} \cdot \frac{36}{37} \cdot \dots \cdot \frac{36}{37},$$

であり、これは $p^3 q^{17}$ と書くのと同じです。これは、3回の成功と17回の失敗に対応する一つの結果を計算するための公式です。しかし、20回の試行の中に3回の成功を配置する方法は $C_{20,3}$ 通りあります。これらの列は互いに排反な結果に対応するので、それらの確率を足し合わせます。その結果、 $C_{20,3} p^3 q^{17}$ が得られます。

$$P(\text{ちょうど3人が勝つ}) = \underbrace{p^3 q^{17} + p^3 q^{17} + \dots + p^3 q^{17}}_{C_{20,3} \text{ 回}} = C_{20,3} \cdot p^3 q^{17},$$

これは前と同じ公式です。

4.2 ルートボックス

あるゲームでは、各プレイでレアなルートアイテムが出現する確率が0.26%です。各プレイには\$2.49かかります。「成功」をルートアイテムを手に入れることとします。

問題

- (a) 100回の試行でちょうど2回勝つ確率はいくらでしょうか。
- (b) 1000回の試行でちょうど3回勝つ確率はいくらでしょうか。
- (c) ルートアイテムを見つけるまでに、いくら払うことを期待すべきでしょうか。

解答

成功確率は

$$p = 0.0026, \quad q = 1 - p = 0.9974.$$

です。

- (a) ここでは $n = 100$, $r = 2$ です。したがって、

$$P(X = 2) = C_{100,2}(0.0026)^2(0.9974)^{98}.$$

数値的には、およそ

$$P(X = 2) \approx 0.0259 = 2.59\%.$$

です。

- (b) ここでは $n = 1000$, $r = 3$ です。したがって、

$$P(X = 3) = C_{1000,3}(0.0026)^3(0.9974)^{997}.$$

数値的には、およそ

$$P(X = 3) \approx 0.218 = 21.8\%.$$

です。

- (c) 期待される勝利数が1に達するような試行回数 n を求めています。 $E(X) = np$ なので、次を解きます。

$$1 = n(0.0026) \quad \implies \quad n = \frac{1}{0.0026} \approx 384.6.$$

したがって、平均的には、ルートアイテム1個に対して約385回プレイする必要がありますと期待できます。各プレイは\$2.49なので、費用は

$$385 \times 2.49 = \$958.65.$$

です。

4.3 回転寿司

ある回転寿司店では、皿を5枚機械に入れるとゲームをして賞品を当てることができます。賞品が当たる確率は25%です。

問題

- (a) 5回ゲームをして、2回勝つ確率はいくらでしょうか。
- (b) 2回の試行で2回勝つ確率はいくらでしょうか。
- (c) 平均的に、賞品を一つ当てるまでに何枚の皿を食べると期待すべきでしょうか。

解答

成功を賞品に当たることと定義します。すると、

$$p = 0.25, \quad q = 0.75.$$

です。

- (a) ここでは $n = 5, r = 2$ なので、

$$P(X = 2) = C_{5,2}(0.25)^2(0.75)^3.$$

数値的には、

$$P(X = 2) = 10(0.0625)(0.421875) = 0.263671875 \approx 26.4\%.$$

です。

- (b) 2回の試行で2回勝つということは、両方の試行で勝つという意味です。試行は独立なので、

$$P(\text{勝ち、その後勝ち}) = 0.25 \times 0.25 = 0.0625 = 6.25\%.$$

- (c)

- (d) $E(X) = np$ なので、 $E(X) = 1$ となるようにしたいです。したがって、

$$1 = n(0.25) \quad \implies \quad n = 4.$$

つまり、一つの賞品に当たるまでに約4回の試行を期待します。各試行には皿5枚が必要なので、これは平均で

$$4 \times 5 = 20$$

枚の皿に対応します。

4.4 サイコロを振る

標準的な6面サイコロを3回振るとします。

「成功」を6を出すこととします。すると、各試行には二つの結果（6または6でない）しかなく、試行は独立であり、成功確率は一定です。

$$n = 3, \quad p = P(6 \text{ を出す}) = \frac{1}{6}, \quad q = \frac{5}{6}.$$

問題

ちょうど2回6が出る確率はいくらでしょうか。

解答

ちょうど2回6を出すには、 $n = 3$ 回の試行のうち $r = 2$ 回の成功が必要です。二項公式より、

$$P(r \text{ 回成功}) = C_{n,r} p^r q^{n-r}.$$

したがって、

$$P(\text{ちょうど2回 } 6) = C_{3,2} \left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right) = 3 \cdot \frac{1}{36} \cdot \frac{5}{6} = \frac{15}{216} = \frac{5}{72} \approx 6.9\%.$$

説明としては、3回のうちの2回が6になるかを選ぶ方法は $C_{3,2} = 3$ 通りあり、そのような各パターンの確率は $\left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right)$ です。

4.5 バスケットボールをする猫

猫がバスケットボールのフリースローをしています。猫はバスケットボールがあまり得意ではないので、シュートが入る確率を $p = 0.5$ とします。この猫は6本シュートを打ちます。

「成功」を「シュートが入ること」とします。すると、

$$n = 6, \quad p = 0.5, \quad q = 0.5.$$

です。

問題

- (a) 猫がちょうど4本シュートを決める確率はいくらでしょうか。
- (b) 猫が少なくとも4本シュートを決める確率はいくらでしょうか。

解答

X を決まったシュートの本数とします。

- (a) ちょうど4本決めるとは、 $n = 6$ 回の試行のうち $r = 4$ 回成功するという意味です。

$$P(X = 4) = C_{6,4} (0.5)^4 (0.5)^2 = C_{6,4} (0.5)^6 = \frac{C_{6,4}}{64} = \frac{15}{64} \approx 23.4\%.$$

- (b) 「少なくとも4本」とは、4本、5本、または6本決めるという意味なので、これらの互いに排反な場合を足し合わせます。

$$\begin{aligned} P(X \geq 4) &= P(X = 4) + P(X = 5) + P(X = 6) \\ &= \frac{C_{6,4}}{64} + \frac{C_{6,5}}{64} + \frac{C_{6,6}}{64} = \frac{15 + 6 + 1}{64} = \frac{22}{64} = \frac{11}{32} \approx 34.4\%. \end{aligned}$$