

1 日目

確率分布の基本



未来のことにように、まだ起きていないできごとが起きる可能性を数学的に測る、という視点が明確に文献に残ったのは、パスカルとフェルマーの手紙が最初とされます。

「2 人が同じ金額を出し合って、先に 3 勝した方が、掛け金を全額もらえるという賭けをしたが、一方が 2 勝 1 敗したところで、やむをえず勝負を中止することになった。このとき掛け金はどう分配すべきか？」

この問題は、パスカルが友人シュバリエ・ド・メレから相談されたもので、パスカルは数学者であるフェルマーに答えを求めたのですが、この手紙をきっかけにパスカルとフェルマーの間で往復書簡が始まり、現代の確率論につながる議論へと発展したと言われています。

1 日目 確率分布の基本

1 確率変数・確率分布って何だ!?

ねらい 確率分布を学ぶ上で基本となる用語や記号を整理しておきましょう。

確率分布を極めるには、とにかくにも「確率変数」って用語を押さえておかねば！
ってことで、まずは基本の用語をまとめておきます。

1 確率変数と確率分布

ものすごく単純に言うと、**確率変数**とは、確率によっていろいろな値をとる変数のことを言います。
例えば…。

コインを2回投げたときに、表が出る回数 X とその確率について考えてみます。

このとき、コインの表裏の組合せとしては、(表, 表), (表, 裏), (裏, 表), (裏, 裏)の4通りがあるので、表が出る回数とその確率は

$$2 \text{ 回 } (X=2 \text{ の場合}) : \frac{1}{4}, \quad 1 \text{ 回 } (X=1 \text{ の場合}) : \frac{2}{4}, \quad 0 \text{ 回 } (X=0 \text{ の場合}) : \frac{1}{4}$$

となりますね。

このように、表の出る回数 X は、確率によって変化するので**確率変数**です！

で、**確率分布**とは、**確率変数**がとり得る値とその確率を対応させたものです。

上の例で言うと

X	0	1	2	計
確率	$\frac{1}{4}$	$\frac{2}{4}$	$\frac{1}{4}$	1

「計」が1になるのはアタリマエだよ！
だって、表が出る回数は0回か1回か2回しかないから、それらの確率を合計すると「必ず起こる確率」になるからね。

これまで変数と呼んでいた x や y は、「1/5の確率で1、それ以外で0」みたいに、確率に左右されることはありませんでしたわ。そのため、**確率変数**という特別なヤツを導入するんです。

のように表にまとめることができます（この表を**確率分布表**と呼びます）。

そして、**確率変数** X はこの**確率分布**に**従う**といいます。

確率分布は、記号を使ってまとめることもできます。上の例では

$$P(X=0)=\frac{1}{4}, \quad P(X=1)=\frac{2}{4}, \quad P(X=2)=\frac{1}{4}$$

ってな具合です。

確率変数がある値をとる確率

確率変数 X のとりうる値が $x_1, x_2, \dots, x_n$ であるとき、 X が1つの値 x_k をとる確率を **$P(X=x_k)$** で表します。

X の値が a 以上 b 以下である確率を **$P(a \leq X \leq b)$** と表すこともできます。

上の例で言えば、

$$P(1 \leq X) = \frac{2}{4} + \frac{1}{4} = \frac{3}{4}$$

「表が出る回数が少なくとも1回である確率」と表現することもできます。

などと表せますヨ。

「確率分布と統計的な推測」の単位では、数Aで学習した「確率」を、もっと統計的な見方で深めようって話になります。確率が苦手な人は多いのですが、この単位は統計に重きが置かれるので、数Bで扱う確率はとても簡単なものになっています。ご安心を（たぶん……）。

2

平均・分散・標準偏差

数Ⅰの「データの分析」で、平均、分散、標準偏差を学習しました。今回学んでいる「確率分布と統計的な推測」でも、同じ用語が登場します。これらは、データの分析で学んだものを、より統計分野で扱いやすくするために拡張したものになっています。← 定義自体は同じですのでご安心ください。

ここでは、確率変数 X が従う確率分布が

X	x_1	x_2	$\cdots$	x_n	計
P	p_1	p_2	$\cdots$	p_n	1

となることを考えます。

平均（期待値）

いきなりですが、定義をどうぞ。

平均（期待値）

$x_1p_1 + x_2p_2 + \cdots + x_np_n = \sum_{k=1}^n x_k p_k$ を、確率変数 X の平均または期待値といい、 $E(X)$ で表します。つまり

$$E(X) = x_1p_1 + x_2p_2 + \cdots + x_np_n = \sum_{k=1}^n x_k p_k$$

例えば、コインを2回投げたときに表が出る回数を X としたとき、その確率分布は

X	0	1	2	計
確率	$\frac{1}{4}$	$\frac{2}{4}$	$\frac{1}{4}$	1

となるので、平均 $E(X)$ は

$$E(X) = 0 \times \frac{1}{4} + 1 \times \frac{2}{4} + 2 \times \frac{1}{4} = 1 \quad \cdots (*)$$

と計算できます！

● COLUMN : 結局、平均って何者？ ●

上の(*)式は

$$E(X) = \frac{0 \times 1 + 1 \times 2 + 2 \times 1}{4}$$

と変形することもできますね。これは、データ「0, 1, 1, 2」の和をその個数4で割ったものと見ることができます。これはまさに4個のデータの平均に他なりません。

つまり、確率変数の平均は、これまで考えていた平均の拡張として自然なんですね。

「期待値」は、「確率変数 X がとる値として最も期待できるもの」ってニュアンスです。

分散

これも定義からいきましょう！

分散

平均が m となる確率変数 X の分散 $V(X)$ を次式で定義します。

$$\begin{aligned} V(X) &= \sum_{k=1}^n (x_k - m)^2 p_k \\ &= (x_1 - m)^2 p_1 + \cdots + (x_n - m)^2 p_n \end{aligned}$$

これは、確率変数 $(X - m)^2$ の平均を求めたことと同じなので、 $V(X) = E((X - m)^2)$ と表すこともできます。意味は後回しにして、とりあえず計算してみましょう。

コインを2回投げたときに表が出る回数を X として、 X の分散を求めてみましょう。そのために、 X から平均（さっきの例から、 $m = E(X) = 1$ ）を引いたもの（**偏差**といいます）と、その2乗を計算しておきましょう。

X	0	1	2	計
$X - m$	-1	0	1	
$(X - m)^2$	1	0	1	
P	$\frac{1}{4}$	$\frac{2}{4}$	$\frac{1}{4}$	1

ここを見れば分散が求められます。

ってなワケで、分散 $V(X)$ は

$$V(X) = 1 \times \frac{1}{4} + 0 \times \frac{2}{4} + 1 \times \frac{1}{4} = \frac{1}{2}$$

と求めることができます。

COLUMN : 分散は何者か？

分散は、一言で言うと「確率変数 X の値が、その平均からどの程度ズれているのか？」を表した量です。そのココロを見ていきましょう。

確率変数 X のとり得る値 $x_1, x_2, \dots, x_n$ が平均 m からどの程度ズれているかを示す量として、その差（偏差）である、 $x_1 - m, x_2 - m, \dots, x_n - m$ が考えられます。

が、確率変数 $X - m$ の平均を計算してみると…

$$\begin{aligned} E(X - m) &= \sum_{k=1}^n (x_k - m)p_k \\ &= (x_1 - m)p_1 + (x_2 - m)p_2 + \dots + (x_n - m)p_n \\ &= (x_1p_1 + x_2p_2 + \dots + x_np_n) - m(p_1 + p_2 + \dots + p_n) \\ &= m - m = 0 \end{aligned}$$

これは平均です。

確率の和は1です。

となって、必ず0になってしまいます。これでは意味がありませんね。

そこで、偏差ではなく、偏差の2乗を考えることによって、確率変数 X の値が平均からどの程度ズれているかを表してみようと考えだされたのが分散なのです。

データの分析で学んだように、分散は

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

で計算することもできます。

さっきの例で言えば

$$E(X^2) = 0^2 \times \frac{1}{4} + 1^2 \times \frac{2}{4} + 2^2 \times \frac{1}{4} = \frac{3}{2}$$

なので

$$V(X) = E(X^2) - \{E(X)\}^2 = \frac{3}{2} - 1^2 = \frac{1}{2}$$

と求めることができます！

分散の定義式を変形することで証明できます。

$$\begin{aligned} V(X) &= \sum_{k=1}^n (x_k - m)^2 p_k = \sum_{k=1}^n (x_k^2 - 2mx_k + m^2) p_k \\ &= \sum_{k=1}^n x_k^2 p_k - 2m \sum_{k=1}^n x_k p_k + m^2 \sum_{k=1}^n p_k \\ &= \sum_{k=1}^n x_k^2 p_k - 2m \cdot m + m^2 \cdot 1 = \sum_{k=1}^n x_k^2 p_k - m^2 \end{aligned}$$

よって

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

標準偏差

分散は、偏差を2乗した関係で、単位が X の2乗になってしまいます。

例えば、 X の単位が m であれば、分散の単位は m^2 になってしまうのです。長さの単位が面積の単位になってしまうので、これはおかしいですね？

そこで、分散の正の平方根をとった $\sqrt{V(X)}$ を考え、これを X の標準偏差といって $\sigma(X)$ で表します。つまり $\sigma(X) = \sqrt{V(X)}$ です！

「シグマ」と読みます。
和の記号 Σ の小文字です。

3

確率変数の変換

確率変数 X が変化したとき、それに伴って平均・分散・標準偏差はどのように変化するのを見ていきましょう。いちいち始めから計算し直す必要はありませんよ。

 X が a 倍になった場合

ある確率変数 X が a 倍、すなわち aX となったとき、次のことが成り立ちます。

確率変数の変換その 1

- ㊦ X を a 倍すると、平均も a 倍になる： $E(aX) = aE(X)$
- ㊦ X を a 倍すると、分散は a^2 倍になる： $V(aX) = a^2V(X)$
- ㊦ X を a 倍すると、標準偏差は $|a|$ 倍になる： $\sigma(aX) = |a|\sigma(X)$

「コインを投げて、表が出た回数をそのまま点数 X とする」ゲームを例にイメージしてみましょう。

もし、「1 回表が出たら 1 点ではなく、2 倍の 2 点をあげよう」というルールに変わったとしたら、得点は $2X$ となります。点数が 2 倍になれば、その平均も 2 倍になるのは当たり前ですね。

また、点数が 2 倍になれば個々の差も 2 倍に広がるので、平均からの隔たりを示す量である標準偏差も 2 倍になります。そうすると、標準偏差を 2 乗した分散は、 $2^2=4$ 倍となるわけです。これがザッとしたイメージです。

 X に定数 b を足した場合

ある確率変数 X に b を足す、すなわち $X+b$ となったとき、次のことが成り立ちます。

確率変数の変換その 2

- ㊦ X に b を足すと、平均も b だけ増える： $E(X+b) = E(X) + b$
- ㊦ X に b を足しても、分散は変わらない： $V(X+b) = V(X)$
- ㊦ X に b を足しても、標準偏差は変わらない： $\sigma(X+b) = \sigma(X)$

さっきのゲームで、「ハンデとして最初から 1 点を与えよう」というルールに変わったとすれば、平均が 1 点増えるのは当たり前だし、全員一律で点数を与えているのだから個々の平均からの隔たりが変わることもありません。したがって、分散も標準偏差も変わりません。そりゃそーです。

以上のことを合わせると、次のようになります。

確率変数の変換のまとめ

確率変数 X と、定数 a, b に対して

$$\begin{aligned} E(aX+b) &= aE(X) + b \\ V(aX+b) &= a^2V(X) \\ \sigma(aX+b) &= |a|\sigma(X) \end{aligned}$$

が成り立ちます！

これらの関係は、丸暗記ではなく、あくまでもイメージとして覚えておきましょう。

4

独立な確率変数とは？

ある2つの確率変数 X と Y が互いに影響しないとき、 X と Y は**独立**であるといいます。この意味を探るために、例を3つほど考えてみましょう。

例1 さいころで1回目に出た目の数を X 、2回目に出た目の数を Y とする。

☞ 1 回目に何が出ようと、2 回目の出目に影響はありませんね。つまり、 X がどんな値であろうが、 Y の値には一切影響ありません。よって、 X と Y は**独立**です。

例2 0, 1, 2, 3 の数字が書かれた4つのくじを引いて、1回目に引いたくじの数字を X 、そのくじを元に戻して2回目に引いたくじの数字を Y とする。復元抽出っていい嘛。

☞ 1 回目の数字 X が 2 回目の数字 Y に影響することはありません。よって、この場合も X と Y は**独立**ですね。

例3 例2と同じ状況で、1回目に引いたくじの数字を X 、そのくじを元に戻さずに2回目に引いたくじの数字を Y とする。非復元抽出っていい嘛。

☞ 例えば、1回目に0を引いた場合、2回目に0を引くことはありません。つまり、2回目の数字が1回目に引いた数字に影響されます。よって、 X と Y は**独立でない**ことになります。

というわけで、確率変数の独立を式で定義しておきましょう。

独立な確率変数

X のとり得る任意の値 a と、 Y のとり得る任意の値 b について

$$P(X=a, Y=b) = P(X=a)P(Y=b)$$

が成り立つとき、確率変数 X と Y は互いに**独立**であるといいます。

上の例2、例3について、それぞれ例として $P(X=1, Y=2)$ と $P(X=1)P(Y=2)$ を計算してみましょう。

例2の場合

$$P(X=1, Y=2) = \frac{1}{4} \times \frac{1}{4} = \frac{1}{16}$$

$$P(X=1) = \frac{1}{4}, \quad P(Y=2) = \frac{1}{4} \quad \text{より} \quad P(X=1)P(Y=2) = \frac{1}{4} \times \frac{1}{4} = \frac{1}{16}$$

となって、 $P(X=1, Y=2) = P(X=1)P(Y=2)$ が成り立ちます。同じように考えると、 $X=1$ 、 $Y=2$ の場合に限らず成り立つことがわかるので、 X と Y は独立であると言えます！

例3の場合

2回目に引くときは、1回目に引いたくじ以外の3つから選ぶことになるので

$$P(X=1, Y=2) = \frac{1}{4} \times \frac{1}{3} = \frac{1}{12}$$

で、1回目は4つから引けるので $P(X=1) = \frac{1}{4}$ ですが、 $P(Y=2)$ は、1回目で2以外のくじを引き、2回目で2のくじを引く場合を考えるので

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

となります。したがって、 $P(X=1, Y=2) \neq P(X=1)P(Y=2)$ となるので、 X と Y は独立ではありません。

確率の分野で、「独立」という用語は、いろいろな場面で出てきますが、ここでは「事象の独立」を考えてみます。……いきなりですが、定義です。

事象の独立

2つの事象 A, B について

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

が成り立つとき、それらの事象 A, B は**独立**であるといいます。また、事象 A, B が独立でないときは、2つの事象は**従属**であるといいます。

例えば、大小2個のさいころを同時に投げるとき

- ① 大きいさいころの目が1であるという事象を A
- ② 小さいさいころの目が2であるという事象を B
- ③ さいころの目の和が6であるという事象を C

とします。

このとき、2つの事象 A, B については

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

より、 $P(A \cap B) = P(A)P(B)$ が成り立ち、 A と B は独立であると言えます。

一方、2つの事象 A, C については

$$P(A) = \frac{1}{6}, \quad P(C) = \frac{5}{36}, \quad P(A \cap C) = \frac{1}{36}$$

より、 $P(A \cap C) \neq P(A)P(C)$ となってしまうので、 A と C は従属です。

大 小	1	2	3	4	5	6
1	●				●	
2	● ●	●	●	● ●	●	●
3	●		●			
4	●	●				
5	● ●					
6	●					

COLUMN : 事象の独立って何言ってるの？

上の例を見ても、事象の独立って何のことやら？って感じではないでしょうか。

一般に、 $P(A \cap B) = P(A)P(B)$ のとき（つまり、事象 A と B が独立のとき）、両辺を $P(A)$ で割ると

$$\frac{P(A \cap B)}{P(A)} = P(B) \quad \dots\dots ①$$

となりますね。ここで①の左辺は、事象 A が起こったときに事象 B が起こる条件付き確率 $P_A(B)$ に他ならないので、①は

$$P_A(B) = P(B) \quad \leftarrow \text{この式を事象の独立の定義としても OK です。}$$

です。この式が主張していることは

「 A が起きるかどうかを考慮しないで考える B が起きる確率 $P(B)$ 」と

「 A が起きたという条件のもとで B が起きる条件付き確率 $P_A(B)$ 」が等しい！！

ってことなので、要は B が起きることは、 A が起きたかどうかに関係ってことになります。

上の例で、大きいさいころの目を X 、小さいさいころの目を Y とすると

$$P(A) = P(X=1), \quad P(B) = P(Y=2), \quad P(A \cap B) = P(X=1, Y=2)$$

と書き換えることができるので

$$P(A \cap B) = P(A)P(B) \quad \text{と} \quad P(X=1, Y=2) = P(X=1)P(Y=2)$$

は本質的に同じことを言っています。

つまり、「確率変数 X と Y が独立」とは、 X のとり得る任意の値 a と、 Y のとり得る任意の値 b について、事象「 $X=a$ 」と事象「 $Y=b$ 」が独立になるということです。

● COLUMN : 独立と排反は違う？ ●

数 A で学習した事象の**排反**と、今回学んだ事象の**独立**は何が違うのでしょうか？

事象の排反とは、2 つの事象が**同時に起こることがない**ことを言います。例えば、さっきの例の事象 A と B は独立でしたが、排反ではありません。だって、大きいさいころの目が 1 という事象と、小さいさいころの目が 2 という事象は同時に起こることがありますよね？

でも、例えば、大きいさいころを 1 回投げるとき、1 の目が出るという事象と、2 の目が出るという事象が同時に起こることはありません。こんなとき、これらの事象は排反である、というわけです。

6

確率変数の和や積

2 つの確率変数 X , Y の和 $X+Y$ や積 XY の平均や分散がどうなるか考えてみましょう。
一気に紹介します。

確率変数の和や積

(1) ある確率変数 X , Y について

$$E(X+Y) = E(X) + E(Y)$$

(2) 互いに独立な確率変数 X , Y について

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

$$V(X+Y) = V(X) + V(Y)$$

(2) は独立じゃないと成り立たないので注意！

証明は教科書に譲るとして、ここではどんなときに役に立つかを紹介します。

例えば、コインを 10 回投げたときに表が出る回数 Y の平均 $E(Y)$ と分散 $V(Y)$ を求める問題を考えてみましょう。

これをいちいち「10 回表が出る確率は……、9 回表が出る確率は……」なんて 1 つ 1 つ求めていくのは大変ですね。

ですが、思い出してください！

p.4, p.5 で、コインを 2 回投げたときに表が出る回数を X としたときの平均 $E(X)$ と分散 $V(X)$ はすでに求めました。

$$E(X) = 1, \quad V(X) = \frac{1}{2}$$

でしたね。

ということで、事象 Y は事象 X を 5 回（それぞれ X_1, X_2, X_3, X_4, X_5 とする）繰り返したものとと同じになります。で、 $X_1, \dots, X_5$ は独立（前の結果が表裏の出方に影響なし）なので

$$\begin{aligned} E(Y) &= E(X_1 + X_2 + X_3 + X_4 + X_5) \\ &= E(X) + E(X) + E(X) + E(X) + E(X) \\ &= 5 \times E(X) = 5 \times 1 = 5 \end{aligned}$$

$$V(Y) = V(X_1 + X_2 + X_3 + X_4 + X_5)$$

$$= 5 \times V(X) = 5 \times \frac{1}{2} = \frac{5}{2}$$

この確認が大事！

ってな感じに求めることができます。

1 日目 確率分布の基本

2 反復試行を思い出せ!! 二項分布

ねらい 試行の結果が2通りしかない場合の確率分布を考えよう。

反復試行の中で、結果が「起こる」か「起こらない」といった、結果が2つしかないような試行を n 回繰り返したときに、事象が起こる回数 X 、事象が起こる確率 p が従う確率分布を二項分布と呼んで、 $B(n, p)$ で表します。 Binomial distribution (二項分布) の頭文字をとっています。

$X = k$ ($k = 0, 1, 2, \dots, n$) となる確率は、反復試行の確率の式から

$$P(X = k) = {}_n C_k p^k (1-p)^{n-k} \quad (k = 0, 1, 2, \dots, n) \quad \text{……①}$$

となります!

この確率分布が二項分布です。

A が何回目に起きるかを決める方法が ${}_n C_k$ 通り。

そのそれぞれのパターンについて、その確率が $p^k(1-p)^{n-k}$ と考える。

ちょっと一言! なぜ二項分布?

二項定理 $(p+q)^n = \sum_{k=0}^n {}_n C_k p^k q^{n-k}$ において、 $q = 1-p$ としたときの形が①に現れるので二項分布って呼ばれています。

1 具体例で理解しようぜ!

例えば、1個のさいころを3回投げるとき、1の目が出る回数を X とすると、 X のとり得る値は

$$X = 0, 1, 2, 3$$

で、 $X = k$ ($k = 0, 1, 2, 3$) となる確率は

$$P(X = k) = {}_3 C_k \left(\frac{1}{6}\right)^k \left(1 - \frac{1}{6}\right)^{3-k}$$

となります。つまり、確率分布は

X	0	1	2	3	計
P	${}_3 C_0 \left(\frac{1}{6}\right)^0 \left(\frac{5}{6}\right)^3$	${}_3 C_1 \left(\frac{1}{6}\right)^1 \left(\frac{5}{6}\right)^2$	${}_3 C_2 \left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right)^1$	${}_3 C_3 \left(\frac{1}{6}\right)^3 \left(\frac{5}{6}\right)^0$	1

で、これが二項分布ってヤツの正体です。

これは、①で $n = 3$ 、 $p = \frac{1}{6}$ としたものなので、 X は二項分布 $B\left(3, \frac{1}{6}\right)$ に従うと表現できます。

もう1つ例を紹介しましょう。

コインを4回投げたときに、表が出る回数を X としましょう。コインを投げたときの結果は「表が出る」か「裏が出る」の2通りなので、 X は二項分布に従います。

試行回数は4、表が出る確率は $\frac{1}{2}$ なので、 X は二項分布 $B\left(4, \frac{1}{2}\right)$ に従います。

つまり

X	0	1	2	3	4	計
P	${}_4 C_0 \left(\frac{1}{2}\right)^0 \left(\frac{1}{2}\right)^4$	${}_4 C_1 \left(\frac{1}{2}\right)^1 \left(\frac{1}{2}\right)^3$	${}_4 C_2 \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^2$	${}_4 C_3 \left(\frac{1}{2}\right)^3 \left(\frac{1}{2}\right)^1$	${}_4 C_4 \left(\frac{1}{2}\right)^4 \left(\frac{1}{2}\right)^0$	1

ですね!

ある確率変数 X が二項分布に従うときの平均、分散、標準偏差を求める公式を一気に紹介！

二項分布の平均・分散・標準偏差

確率変数 X が二項分布 $B(n, p)$ に従うとき、 X の平均 $E(X)$ 、分散 $V(X)$ 、標準偏差 $\sigma(X)$ は

$$E(X) = np$$

$$V(X) = np(1-p)$$

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

で求められます！

証明の前に、軽くイメージしておきましょう。

平均 $E(X)$ は、 n 回のうち事象が何回くらい起きるかってことでした。1 回ごとに p の確率で起こるので、 n 回起こったら、だいたい np 回くらい起こるってのは直観的にもわかりやすいです。

分散は、平均 $E(X)$ に、起こらない確率 $1-p$ をかければ求められる！！って覚えるのととてもラクですよ。……では、証明してみましょう。

X が二項分布 $B(n, p)$ に従うとは、 $P(X=k) = {}_nC_k p^k (1-p)^{n-k}$ ($k=0, 1, 2, \dots, n$) ということなので

○ 1 回ごとに事象 A が起きる確率が p である反復試行を n 回行う

○ A が起きる回数が X 回

と考えて OK です。このとき、確率変数 X_i ($i=1, 2, 3, \dots, n$) を

$$X_i = \begin{cases} 1 & (i \text{ 回目に } A \text{ が起きた場合}) \\ 0 & (i \text{ 回目に } A \text{ が起きない場合}) \end{cases}$$

と定めると

$$X = X_1 + X_2 + X_3 + \dots + X_n \quad \dots\dots ①$$

となります。

X_i ($i=1, 2, 3, \dots, n$) の確率分布は次の通りです。

X_i	1	0	計
P	p	$1-p$	1

よって

$$E(X_i) = 1 \cdot p + 0 \cdot (1-p) = p \quad \dots\dots ②$$

$$\begin{aligned} V(X_i) &= E(X_i^2) - \{E(X_i)\}^2 \\ &= 1^2 \cdot p + 0^2 \cdot (1-p) - p^2 \\ &= p(1-p) \quad \dots\dots ③ \end{aligned}$$

①より

$$\begin{aligned} E(X) &= E(X_1 + X_2 + \dots + X_n) \\ &= E(X_1) + E(X_2) + \dots + E(X_n) \\ &= p + p + \dots + p = np \end{aligned}$$

← n 個の p の和です。

反復試行を考えているので、確率変数 $X_1, \dots, X_n$ は互いに独立となるから

$$\begin{aligned} V(X) &= V(X_1 + X_2 + \dots + X_n) \\ &= V(X_1) + V(X_2) + \dots + V(X_n) \\ &= p(1-p) + p(1-p) + \dots + p(1-p) = np(1-p) \end{aligned}$$

← n 個の $1-p$ の和です。

よって $\sigma(X) = \sqrt{V(X)} = \sqrt{np(1-p)}$ です！！