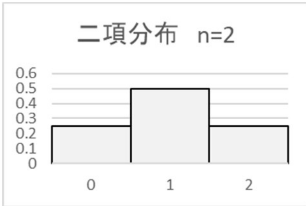
年　組　番　氏名

ＳＳ理数探究Ⅰ　統計処理③（正規分布）

例１　２枚の硬貨を同時に投げるとき、表と裏の出方には

（表，表），（表，裏），（裏，表），（裏，裏）　　の４通りの場合がある。

この試行において表が出る硬貨の枚数をＸとすると，Ｘのとりうる値は，０，１，２であり、Ｘがこれらの各値をとる確率は、次の表のようになる。



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

このＸのように、試行の結果によってその値が定まり、各値に対応して確率が定まるような変数を確率変数という。

一般に、確率変数Ｘのとりうる値が x_1，x_2，・・・，x_n であり、それぞれの値をとる確率が p_1，p_2，・・・，p_n　であるとき、次が成り立つ。

X	x_1	x_2	・・・	x_n	計
P	p_1	p_2	・・・	p_n	1

$p_1 \geqq 0$，$p_2 \geqq 0$，・・・，$p_n \geqq 0$　　　$p_1+p_2+\cdots+p_n=1$

この表の対応関係をＸの**確率分布**といい、確率変数Ｘは**この分布に従う**という。

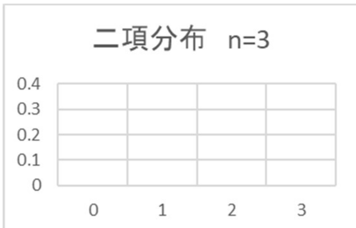
確率変数Ｘが値 a をとる確率を $P(X=a)$ で表す。

また、Ｘが a 以上 b 以下の値をとる確率を $P(a \leqq X \leqq b)$ で表す。

例１において　$P(X=0)=\frac{1}{4}$　，$P(1 \leqq X \leqq 2)=\frac{2}{4}+\frac{1}{4}=\frac{3}{4}$　　となる。

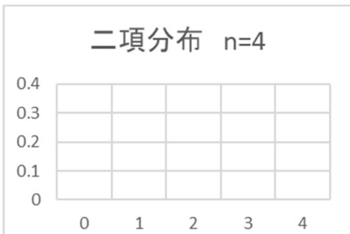
練習１　３枚の硬貨を同時に投げるとき、表が出る硬貨の枚数Ｘの確率分布を求めよ。
また、そのヒストグラムをかけ。

X	0	1	2	3	計
確率Ｐ					



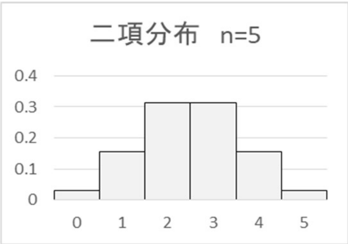
練習２　４枚の硬貨を同時に投げるとき、表が出る硬貨の枚数Ｘの確率分布を求めよ。
また、そのヒストグラムをかけ。

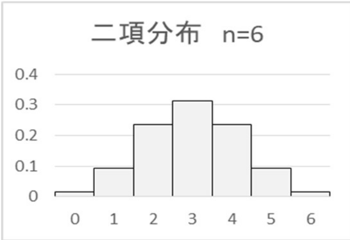
X	0	1	2	3	4	計
確率Ｐ						

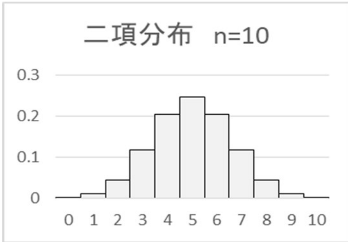


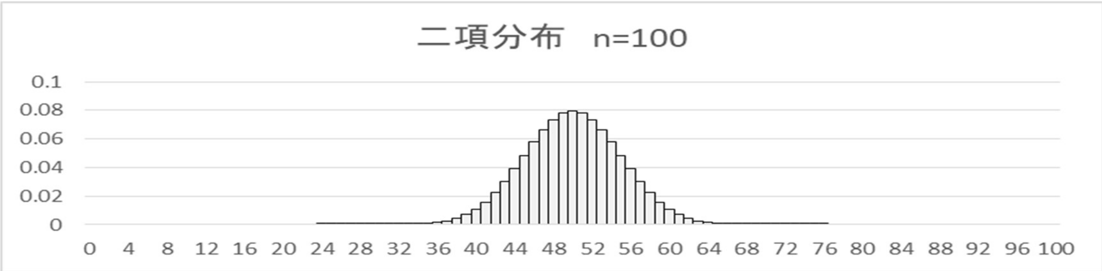
このような分布を**二項分布**といいます。

ここで、硬貨の枚数ｎを増やしていくとどのようなヒストグラムになると思いますか。
硬貨の枚数ｎをｎ＝5，6，10，100 とすると、二項分布はつぎのようになります。

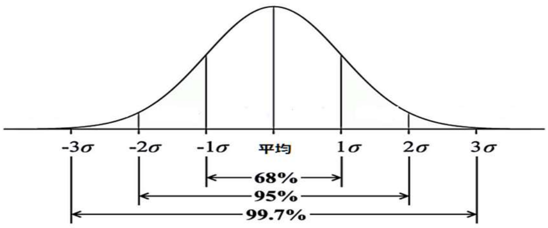








これらの棒グラフの柱の高さの合計は 1 になっています。
この各柱の底辺の長さを 1 とすると、柱の面積の合計も 1 となります
硬貨の枚数を際限なく大きくしていくとどうなるでしょうか。
　次の図のように、左右対称なつりがね状の曲線に近づいていきます。

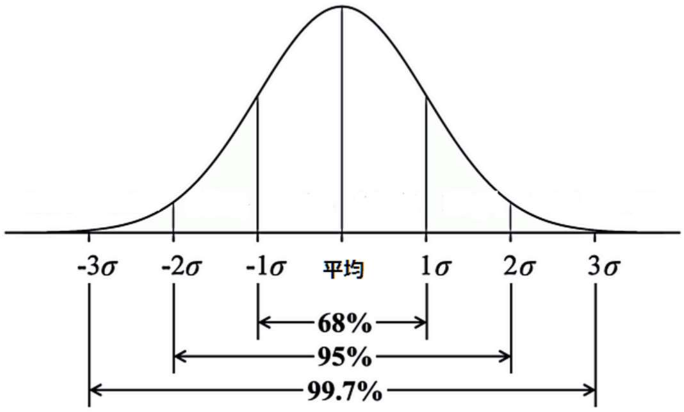


この曲線と横軸とで囲まれた図形の面積は 1 となります。
この曲線を**正規分布曲線**といいます。
正規分布曲線は、ｍを実数，σを正の実数とすると

関数　$f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-m)^2}{2\sigma^2}}$　で表されます。（この式を覚える必要はありません）

ここで，e は無理数の定数で，$e=2.71828\cdots$ となります。
確率変数 X の分布がこの曲線で表されるとき、**確率変数 X は正規分布 N（m，σ^2）に従う**といいます。このとき、Xの平均値はｍ，標準偏差はσとなり、確率について
P(ｍ－σ≦X≦ｍ＋σ)＝0.6827　　　P(ｍ－2σ≦X≦ｍ+2σ)＝0.9545
P(ｍ－3σ≦X≦ｍ+3σ)＝0.9973　　であることが知られています。
特に、標準偏差σの大小に関係なく、ｍ―σからｍ＋σ　の間の面積は同じになります。

統計処理③



★正規分布では、ｘと平均値ｍとの差　ｘ－ｍ　が標準偏差σの何倍であるかがわかれば、データ全体の中でのｘの位置を知ることができます。

・正規分布は実用性が高く、毎年の降雨量、同性同年代の人の身長、人間の知能や**テストの得点**など、自然現象や社会現象には正規分布しているものが多くあります。

○個々のデータの位置を知る

問１　次の表はある５人の生徒の国語と数学の得点です。

No.	1	2	3	4	5	平均点
国語	90	80	70	60	50	70
数学	40	90	30	20	10	38

No.1 の生徒の国語と No.2 の生徒の数学は、どちらも 90 点で同じです。
「No.1 の国語の成績と No.2 の数学の成績は同じである」といってよいでしょうか。

平均点が異なるので、得点と平均点との差を考えてみましょう。

No.1 の国語の得点と平均点の差は：

No.2 の数学の得点と平均点の差は：

よって No.（　　　）の（　　　　）の成績のほうがよいということがわかります。

問２　次の表はある 10 人の生徒の国語と数学の得点です。

No.	1	2	3	4	5	6	7	8	9	10	平均点	標準偏差
国語	90	57	56	54	53	52	50	45	40	33	53	14.3
数学	93	90	80	63	55	45	40	27	20	17	53	26.6

統計処理③

No.1 の国語と No.2 の数学は、どちらも 90 点で同じです。平均点も 53 点で同じです。
「No.1 の国語の成績と No.2 の数学の成績は同じである」といってよいでしょうか。

得点と平均点との差はどちらも（　　　　）点で同じです。
何を比べればよいでしょうか。

データのばらつき具合をくらべてみましょう。
標準偏差を見ると、国語は（　　　　）点、数学は（　　　　）点です。
よって、(　　　　）のほうが得点のばらつきが小さく、よい点を取りにくいけれども
悪い点を取ることも少ないことがわかります。　すなわち、(　　　　）の 90 点のほうが、
（　　　　）の 90 点よりも取りにくく、相対的な価値が高いことになります。
よって No.（　　　）の（　　　　）の成績のほうがよいということがわかります。

問３　次の表はある学級全員の平均点と標準偏差と、２人の生徒の国語と数学の得点です。

学級	平均点	標準偏差
国語	50	20.0
数学	60	10.0

No.	1	2
国語	90	70
数学	70	90

No.1 の生徒の国語と No.2 の生徒の数学は、どちらも 90 点で同じです。
「No.1 の国語の成績と No.2 の数学の成績は同じである」といってよいでしょうか。

平均点も標準偏差も異なるので、何を比べればよいでしょうか。

（　　　　）と（　　　　　）の差が（　　　　　　）の何倍であるかを示す値(基準値)を
考えてみましょう。　基準値＝

No.1 の国語の基準値は：

No.2 の数学の基準値は：

よって、基準値の高い No.（　　　）の（　　　　）の成績のほうがよいということがわかります。