

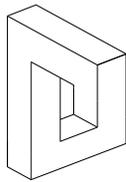
第1章

第2章

第3章

第4章

第5章



第1章 記述統計学

Koishikawa Philosophy II



データやグラフに騙されるな！

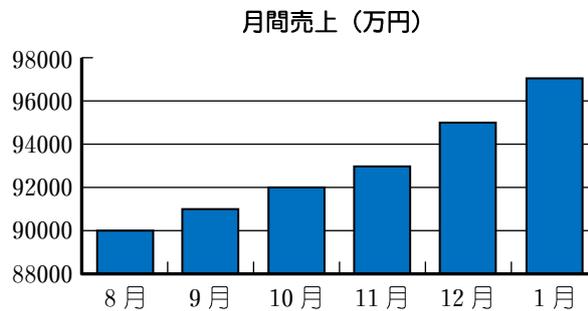
データやグラフの「見せ方」によって、どのように印象が変わるだろう？

統計学は、大量のデータを扱うのにとても便利ですが、その扱い方を誤るととんでもない誤解を生んだり、受ける印象が異なったりしてしまいます。

次の7つのテーマを考え、統計学に潜む落とし穴の怖さを体感してみましょう。

テーマ1 グラフを正しく読み取ろう！～棒グラフ編～

下の棒グラフは、あるスーパーマーケットにおけるあるお菓子の売り上げを、ここ6か月分調査したものです。



このグラフを見て、このお菓子メーカーの担当者は

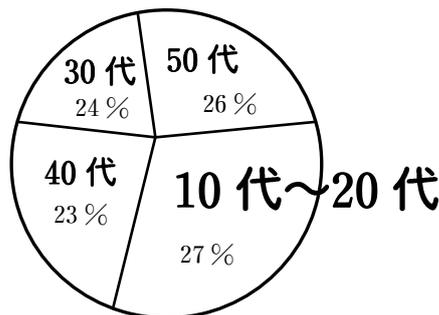
「このお菓子の人気はこの半年で急上昇している！」

と言っています。

この主張に対する反論を考えてください。

テーマ2 グラフを正しく読み取ろう！～円グラフ編～

ある化粧品メーカーが、新発売の化粧品について「購買者の年齢層」を示した下のような円グラフをつくりました。



購買者の年齢層

このグラフをもとにして

「この商品は10代～20代の若者に大人気です」

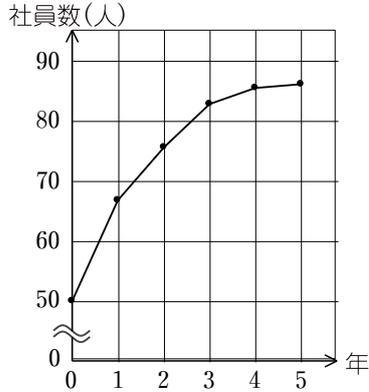
と言っています。

この主張に対する反論を考えてください（このグラフは、とある報道番組で実際に使われたものをアレンジしたものです）。

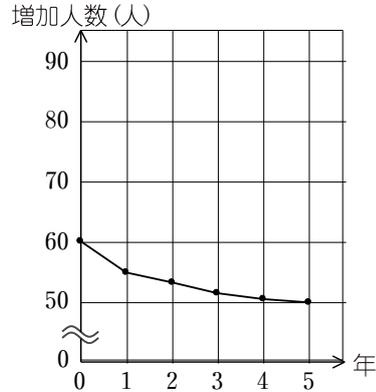
テーマ3 グラフを正しく読み取ろう！～折れ線グラフ編～

下の2つのグラフは、A社、B社のここ5年間の社員数の変化を示したグラフです。

<A社>



<B社>



このグラフを見て、A社の社長は

「わが社の社員数はB社と比べて順調に伸びている！」

と言っています。

この主張に対する反論を考えてください。

テーマ4 データの扱いに気をつけよう！①

あるサプリメントのメーカーが、広告に下のような宣伝をしています。



このメーカーは

「わが社のサプリメントは、信頼と実績があるのでぜひ試してみてください！」

と言っています。

この主張に対する反論を考えてください。

テーマ5 データの扱いに気をつけよう！②

生徒数 830 人の学校で、今年度、インフルエンザにかかった生徒は 43 人もいました。その 43 人を調べたところ、約半数の 22 人が予防注射を受けていて、残りの 21 人は受けていないことがわかりました。

予防注射を受けた人の方が、ほんの少しですが、受けなかった人より多いという結果が出てしまいました。この結果を見て、ある人が

「今年度は予防注射が効かなかったんだねえ」

と言っています。

この主張に対する反論を考えてください。

- 第1章
- 第2章
- 第3章
- 第4章
- 第5章

テーマ6 データの扱いに気をつけよう!③

ある中学校の2年生20人が2つのグループに分かれて、数学の試験に挑戦しました。それぞれのグループの点数は次の通りです。

グループ① 73, 85, 86, 72, 84, 96, 75, 66, 73, 85

グループ② 82, 85, 81, 72, 84, 96, 90, 11, 85, 92

この結果を見て、ある人が

「グループ①の方が平均点が高いから、グループ①の方が数学の力がある」

と言っています。

この主張に対する反論を考えてください。

テーマ7 データの扱いに気をつけよう!④

昔、ある国で貨物自動車を、年間台数にして1000台増産する計画を立て、5年後に、生産台数にしてわずかに200台しか増産できませんでした。ところが、その国の発表を見ると

「はじめの目標の80%を達成することができました」

とあります。

1000台のところを200台しか増産できなかったのならば、達成率は

$$200 \div 1000 = 0.2 = 20\%$$

のはずです。

国の発表は正しかったとして、達成率80%の謎を説明してみてください。

班で選んだテーマ：

話し合いメモ

◎ 各班の発表

班	テーマ	発表内容
1班		
2班		
3班		
4班		
5班		
6班		
7班		

第1章

第2章

第3章

第4章

第5章



PPDACサイクルの実践

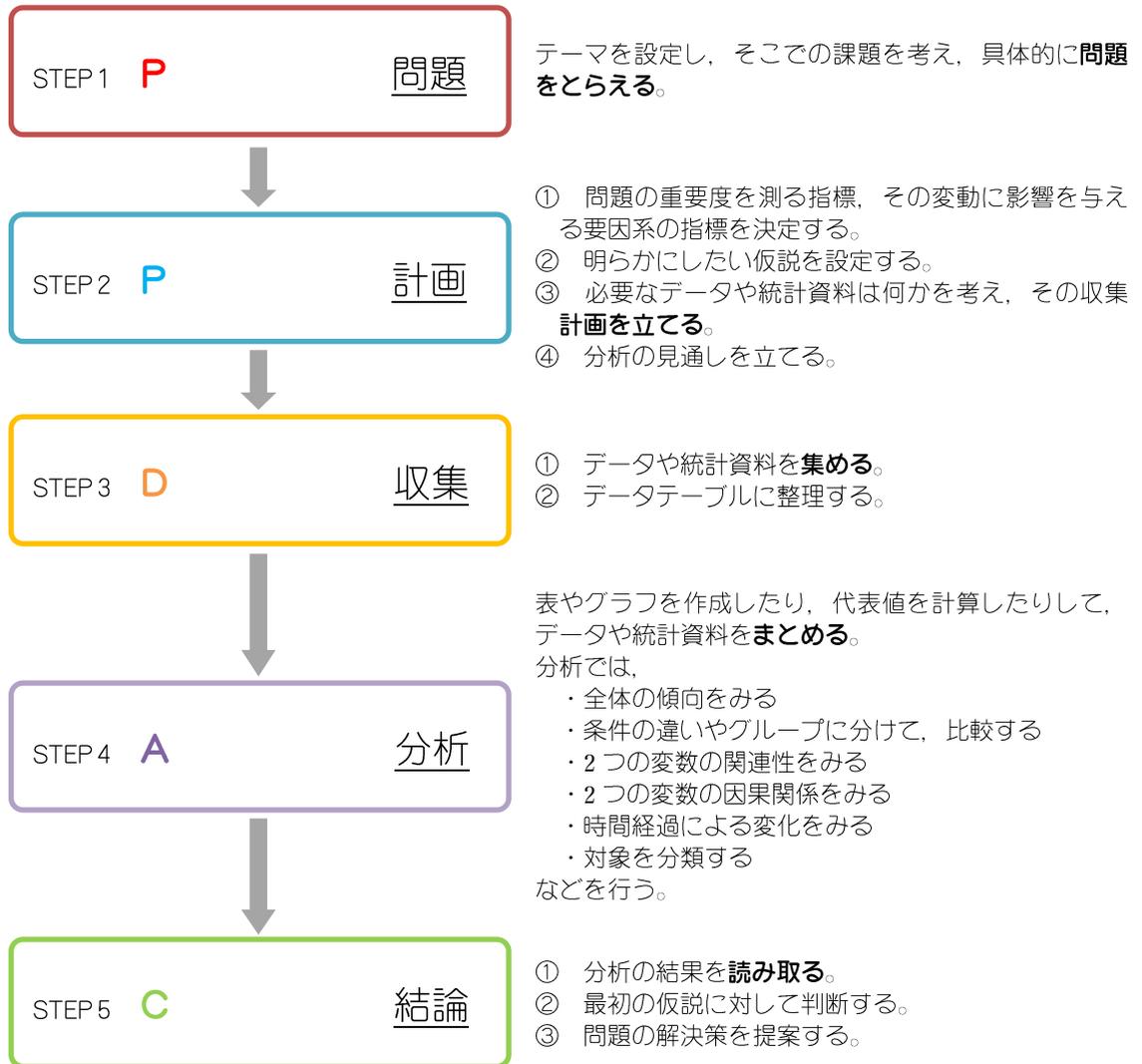
統計的な探究の手法を簡単な例で実践してみよう。

統計的な探究は、以下のプロセスで行われるのが一般的です。

- ① 探究**テーマ**を設定する
- ② 課題を考え、問題を**とらえる**
- ③ 仮説の検証に必要なデータや統計資料を考え、探究**計画を立てる**
- ④ 集団や現象の特徴を知るための統計資料やデータを**集める**
- ⑤ 統計資料やデータを集計・分析し、表やグラフに**まとめる**
- ⑥ 集団や現象の特徴や法則性を**読み取る**
- ⑦ 読み取ったことを、問題の解決策に**活かす**

PPDAC サイクルとは、これらのプロセスを国際的な枠組みとしてまとめたもので、統計的探究の各ステップの頭文字をとって呼んでいます。

PPDAC サイクルは、以下の通りです。



では、PPDAC サイクルの実践として、「**中学生の運動能力の改善**」を考えてみましょう。

第1章

第2章

第3章

第4章

第5章

テレビや新聞で現在の子どもの体力や運動能力が以前より低下していることが取り上げられています。みなさんの周りではどうでしょうか。もし低下していると感じたら、どうして低下してきたと思いますか。実は関係する調査を調べてみると確かに低下している項目もありますが、低下していない項目もあります。もし低下した理由がわかれば、低下が顕著な体力、運動能力について、強化目標を立てて練習するなど、対策がとれるかもしれません。まずは、事実を認識するために、「**昔に比べて現在の子どもの運動能力は低下している**」という仮説を立てて、これを中学生女子のデータについて検証することとしました。

STEP 1

知りたい問題を決めよう

Problem 問題

本当に体力や運動能力が低下しているか、ということ議論する前に、そもそも
(実感として) 体力や運動能力が低下していると感じるか

ということを考えなければなりません。探究しようとしている人が体力の低下を感じていないのならば、この探究には意味がありません。

ここでは、本当に体力や運動能力が低下しているのではないかと感じているという前提で、探究を進めていきます。

グループで、「低下したと思う運動能力は何か」について、話し合しましょう。

<運動能力の例>

握力、上体起こし、長座体前屈、反復横跳び、持久走 1500 m、20 m シャトルラン、50 m 走、立ち幅跳び、ハンドボール投げ など

STEP 2

どのようなデータ・統計資料を集めるか考えよう

Plan 計画

全国ベースの運動能力の実態を知るために、文部科学省が行っている「全国体力・運動能力、運動習慣等調査」から都道府県別の運動能力のデータを調べてみましょう。

タブレットを使って

全国体力・運動能力 運動習慣等調査



で検索してみてください。

また、体力や運動能力について、自分なりに仮説を立ててみましょう。

STEP 3

必要なデータ・統計資料を集めよう

Data 収集

文部科学省「全国体力・運動能力、運動習慣等調査」の都道府県別の結果を見て、上の<運動能力の例>の中から好きなものを選んで、中学生女子の平成 20 年度のデータと令和 4 年度のデータを次のページの表にまとめてみましょう。

中学生女子・ の記録

(単位：)

都道府県	平成 20 年度	令和 4 年度	都道府県	平成 20 年度	令和 4 年度
北海道			滋賀県		
青森県			京都府		
岩手県			大阪府		
宮城県			兵庫県		
秋田県			奈良県		
山形県			和歌山県		
福島県			鳥取県		
茨城県			島根県		
栃木県			岡山県		
群馬県			広島県		
埼玉県			山口県		
千葉県			徳島県		
東京都			香川県		
神奈川県			愛媛県		
新潟県			高知県		
富山県			福岡県		
石川県			佐賀県		
福井県			長崎県		
山梨県			熊本県		
長野県			大分県		
岐阜県			宮崎県		
静岡県			鹿児島県		
愛知県			沖縄県		
三重県					

Analysis 分析

第1章

第2章

第3章

第4章

第5章

(1) いろいろな統計量を計算してみよう

① ……最大値から最小値をひいた値

・平成20年度 _____

・令和4年度 _____

② ……すべてのデータを合計して、データの個数で割った値

・平成20年度 _____

・令和4年度 _____

③ ……データを大きさの順に並べたとき4等分する位置にくる値

	第1四分位数	第2四分位数 (中央値)	第3四分位数
平成20年度			
令和4年度			

(2) 箱ひげ図をかいてみよう

選んだ「運動能力」について、箱ひげ図をかいてみましょう。横軸は調べたデータに合わせて適切に決めてください。



<箱ひげ図から読み取れること>

(3) 散布図をつくってみよう

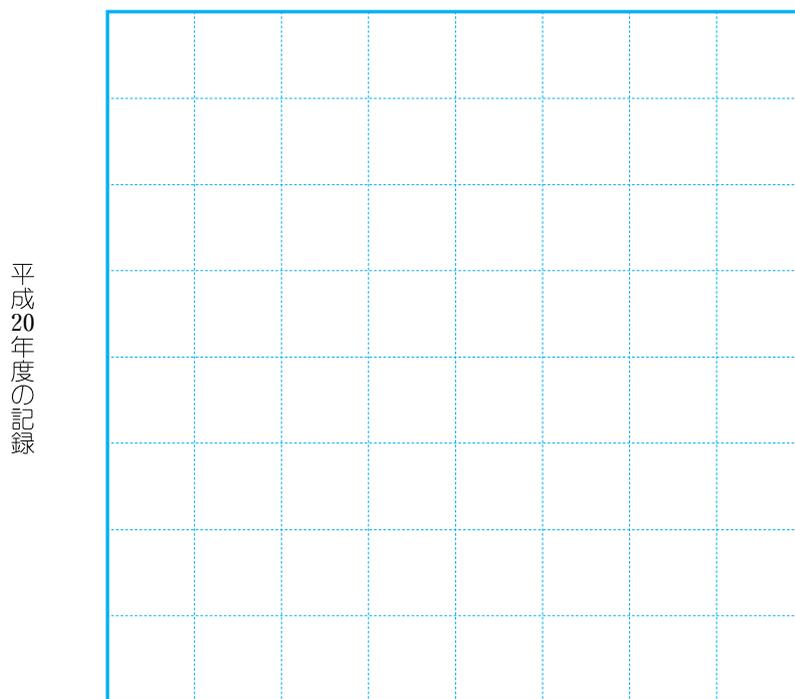
下の図の横軸を平成 20 年度の記録、縦軸を令和 4 年度の記録として

(平成 20 年度の記録、令和 4 年度の記録)

を座標とする点を 47 都道府県についてプロットしてみましょう。

横軸、縦軸の目盛りは記録に合わせて適切に決めてください。

(単位：)

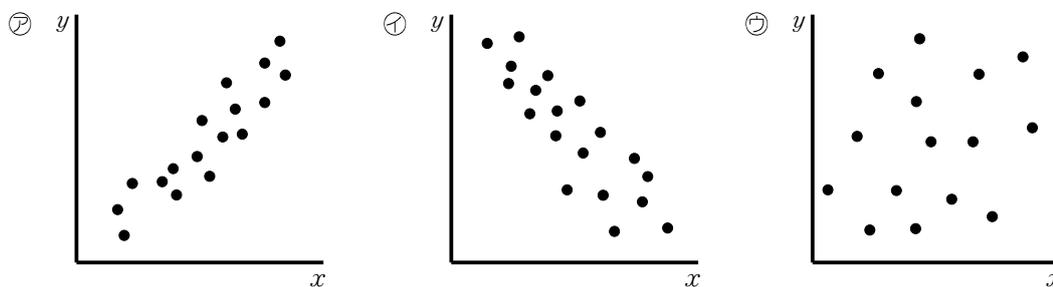


(単位：)

令和 4 年度の記録

上の図のように、2 種類のデータのそれぞれを x 座標、 y 座標として平面上に点をとった図を**散布図**といいます。

散布図の上にどのように点が並ぶかによって、2 つの量（上の例では記録）の関係をざっくり見ることができます。



㊷のように、一方が増えると他方も増える傾向が認められるとき、2 つのデータの間**に正の相関関係がある**といい、㊸のように一方が増えると他方が減る傾向が認められるとき、2 つのデータの間**に負の相関関係がある**といいます。㊹のように、どちらの傾向も認められない場合は、**相関関係はない**といいます。☞ p.16 Column 1. 相関係数

<散布図からわかること>

平成20年度のデータと令和4年度のデータの差が大きい都道府県は、散布図のどのあたりにあるでしょうか。また、日本地図を見ながら、何らかの傾向がないか、結果からわかることを考えてみましょう。

第1章

第2章

第3章

第4章

第5章

STEP5

わかったことをまとめ・読み取ろう

Conclusion 結論

自分が選んだ運動能力の記録について、代表値やグラフ、箱ひげ図、散布図からわかったことをまとめてみましょう。

グラフや図のどの部分に着目して判断したことなのか、その根拠となることを明確にして書きましょう。

◎ この続きは……？

PPDAC サイクルを1回まわしても、目的に合った結論が得られない場合、さらに視点を変えて別の問題（Problem）を考えることにより、PPDAC サイクルをもう1回（あるいはもう2回）まわすこともあります。

今回のテーマでは、例えば

- ・他の運動能力についても調べてみる
- ・男女間で差があるかないかを調べてみる

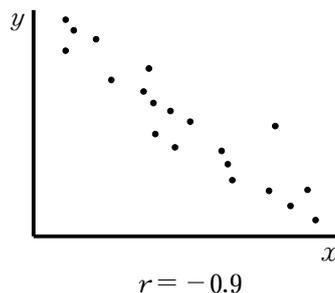
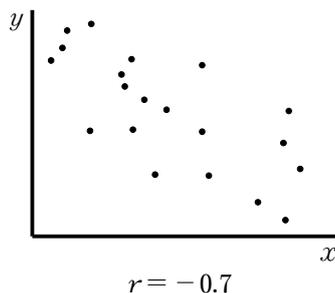
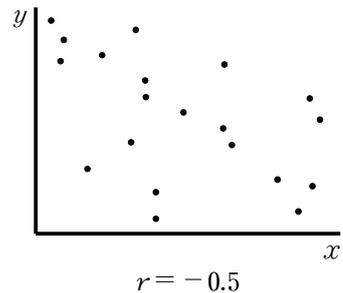
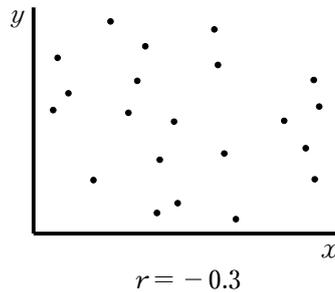
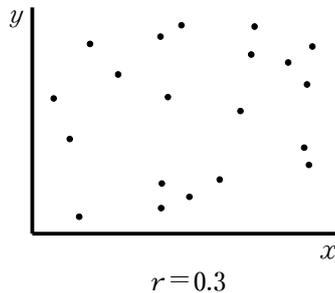
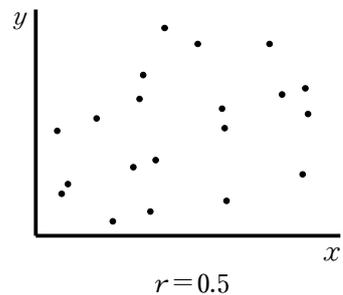
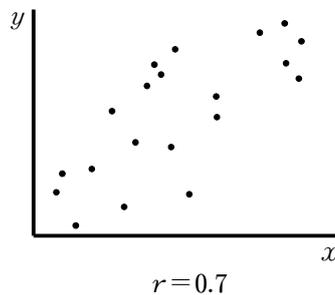
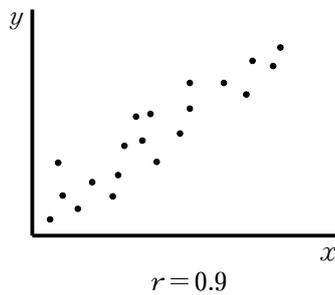
などの調査を考えることができるでしょう。

Column 1. 相関係数

p.14 で、2つのデータの相関関係について軽く触れました。

相関の強さを測る指標として、**相関係数**というものがあります。相関係数は一般的に文字 r で表され、次のような性質があります。

- I. r は -1 から 1 までの値をとる。つまり、 $-1 \leq r \leq 1$ である。
- II. r の値が 1 に近ければ近いほど、**強い正の相関関係**がある。このとき、散布図の点は右上がりの直線に沿って分布する傾向が強くなる。
- III. r の値が -1 に近ければ近いほど、**強い負の相関関係**がある。このとき、散布図の点は右下がりの直線に沿って分布する傾向が強くなる。
- IV. r の値が 0 に近いとき、直線的な相関関係はない。



相関係数は手計算で求めようとするとかなり大変です。高校で学ぶ記号を使って書くと

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}$$

となります (x_i やら \bar{x} の記号の意味はともかくとして、これを計算するのが大変なのはやらなくてもわかりますね)。そこで通常は、表計算ソフトなどを使って計算します。

2. 相関を扱うときの注意～相関関係と因果関係の違い～

相関関係と似た言葉に**因果関係**というものがああります。この違いをきちんとおさえておかないと、データの分析で思わぬミスをしてしまう危険性があります。

相関関係とは、**2 つのことがらのうち、一方が変わると、他方も変わる**という関係です。一方が増えたとき他方も増えるのであれば「正の相関」、一方が増えるとき他方が減るのであれば「負の相関」、どちらでもなければ「相関がない」と呼ばれるのでした。

これに対して因果関係とは、2 つのことがらの一方が**原因**で、他方が**結果**となる関係をいいます。つまり、「A が原因となって、B が起こった」という関係です。

例1 「モテ度」と「バレンタインデーにもらったチョコの数」との関係

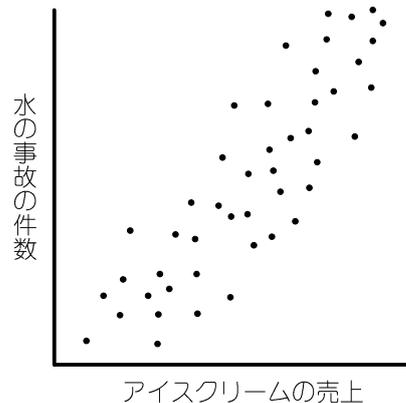
「モテ度」をどのように測るのかはいったんおいておくとして、モテ度が高い人は、より多くのチョコをもらう傾向にあり、相関関係があると言えます。また、モテ度が高いから、もらったチョコが多いという因果関係も成り立つでしょう。

例2 「アイスクリームの売り上げ」と「水の事故の件数」との関係

ある調査の結果、これらの間に右の図のような相関関係が見られたとします。ここでよく勘違いするのが、『**アイスクリームがよく売れる日は水の事故が起きやすいから気をつけよう**』と考えてしまうことです。

アイスクリームの売り上げと水の事故の件数の間には、何の因果関係もありません。

このように、因果関係がない相関関係のことを**疑似相関**（または**偽相関**）といいます。



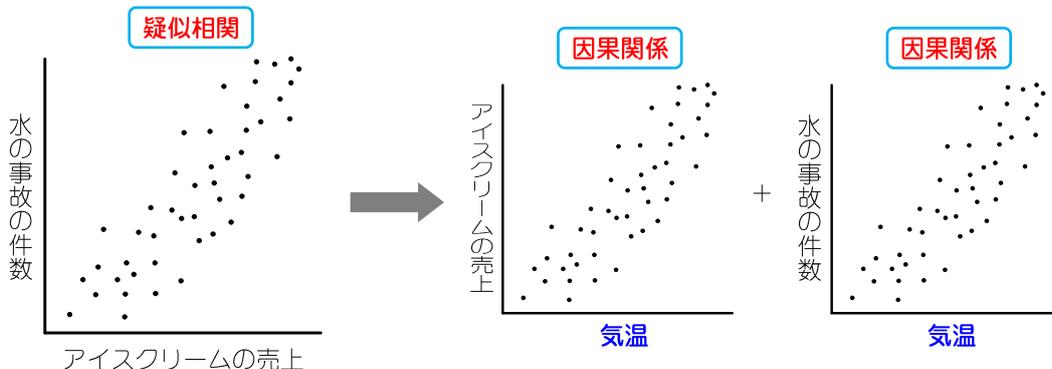
疑似相関がある場合、その要因として考えられるのは次の2パターンです。

- ① まったくの偶然
- ② 別の量との間に因果関係がある場合

②を、上のアイスクリームの売り上げと水の事故の件数の例で考えてみましょう。

アイスクリームの売り上げが増える原因として、例えば気温の上昇が考えられます。また、水の事故の原因も気温の上昇が考えられます（気温が高くなれば水遊びをすることも多くなるので、自然と事故の件数も増えます。夏になると水難事故が起きやすいことからわかるでしょう）。

つまり、下の図のようなイメージです。



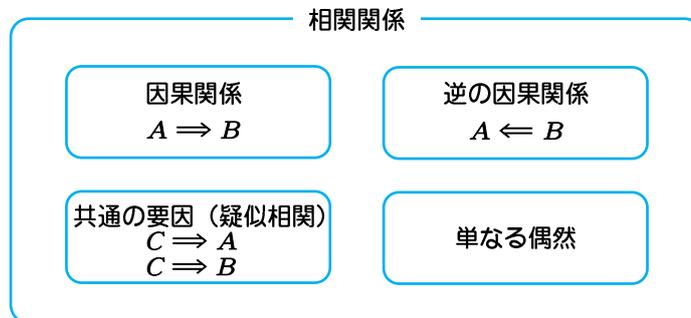
- 第1章
- 第2章
- 第3章
- 第4章
- 第5章

練習 次の(1), (2)はいずれも疑似相関の例です。第3の原因は何だと考えられますか。

(1) 靴のサイズが大きい子どもは文章の読解能力が高い。だから、足を見ればその子の読解能力がわかります。

(2) カラオケ店が多い街ほど犯罪の件数が多い。この街にもう1つカラオケ店を作ったら、犯罪件数がさらに増えるだろう。

2つの量 A , B に相関関係があっても、 $A \Rightarrow B$ という因果関係があるとは限りません。これまでのことをまとめると、次の図のようになります。



ここで、「逆の因果関係」の例を見てみましょう。

例3 「交番の数が多い」(A) 地域ほど、「犯罪件数が多い」(B)

地域に交番が多いほど、それだけ犯罪の抑止力が大きくなるので、犯罪件数は少なくなるはずですが。ところが、実際に調べてみると『交番の数が多い地域ほど犯罪件数も多い』という相関関係が見つかることがあるのです。

この場合、「交番の数が多いから犯罪件数が多い」という因果関係 ($A \Rightarrow B$) は成り立ちません。逆に、「犯罪件数が多い地域だから、交番が多く設置された」という因果関係 ($A \Leftarrow B$) が成り立っています。

上のような相関関係が見つかったからといって、交番の数を減らしてしまえば、逆に犯罪件数が増えるという大失敗につながってしまいます。

3. その他の統計量

この章では、データの特徴をつかむための統計量として、平均値や中央値、四分位数などの代表値を扱いました。ここでは、データの傾向や性質をより精密に分析するために、これまで学んだもの以外にどのような統計量があるかを探っていきます。

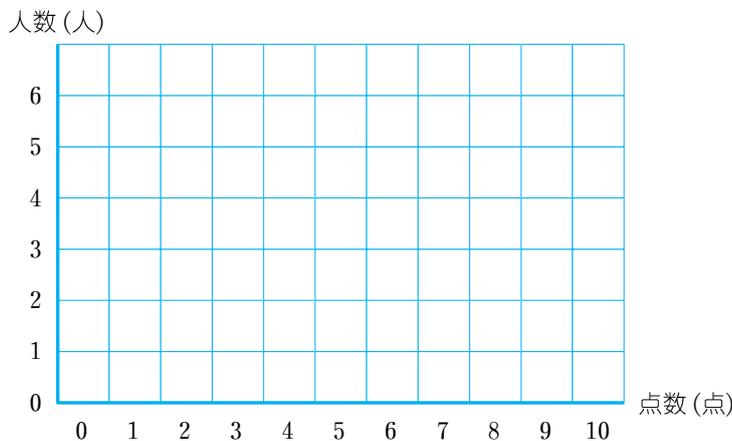
まずは、こんな問題を考えてみましょう。1年生の代数の復習です。

問題 1

10人の生徒が10点満点のテストを受けて、次の表のような結果となったとします。

点数(点)	0	1	2	3	4	5	6	7	8	9	10
人数(人)	3	3	1	0	0	0	0	0	0	0	3

(1) この結果を、下の図にヒストグラムとして表してください。ただし、ここでは階級の幅を0点と考えるので見た目は棒グラフのようになります。



(2) このデータの代表値としてふさわしいものは、平均値、中央値、最頻値のいずれでしょうか。それぞれ求めた上で考えましょう。

- ① 平均値…… 点
- ② 中央値…… 点
- ③ 最頻値…… 点

このデータの代表値としてふさわしいものは、 であると考えます。

◎ その理由は…?

第1章

第2章

第3章

第4章

第5章

ここからが本題。問題1を踏まえて、次の問題2を考えてください。

問題2

A組, B組で, それぞれ10人ずつの生徒が, 10点満点のテストを受けて, 下の表のような結果となったとします。

<A組>

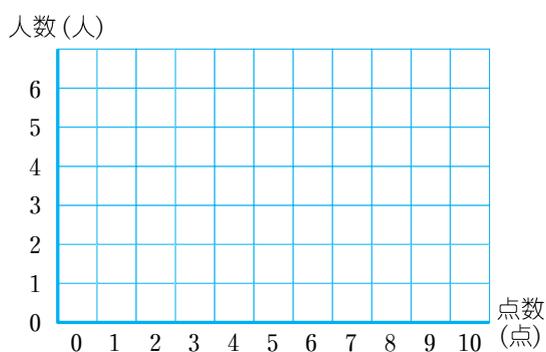
点数(点)	0	1	2	3	4	5	6	7	8	9	10
人数(人)	1	0	0	0	2	4	2	0	0	0	1

<B組>

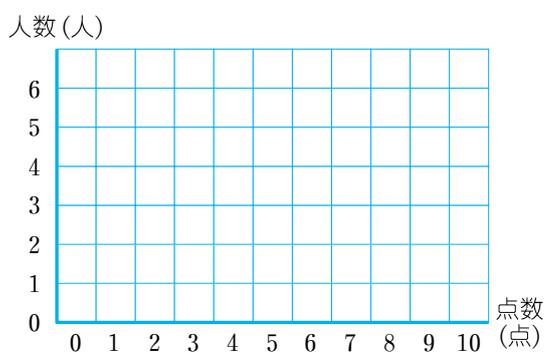
点数(点)	0	1	2	3	4	5	6	7	8	9	10
人数(人)	2	1	0	0	0	4	0	0	0	1	2

(1) それぞれのクラスの手istogramをつくってください。

<A組>



<B組>



(2) それぞれのクラスにおける平均値, 中央値, 最頻値を求めてください。ついでに範囲も求めてみましょう。

<A組>

- 平均値…… 点
- 中央値…… 点
- 最頻値…… 点
- 範囲…… 点

<B組>

- 平均値…… 点
- 中央値…… 点
- 最頻値…… 点
- 範囲…… 点

(3) この2つのデータが区別できるような統計量を考えたいと思います。そこで, 2つのデータからつくった手istogramを見て, 特徴の違いを挙げてみてください。

(4) (3)で調べたことを数値化するために、「個々の点数が平均値（平均点）からどのくらい離れているか」を調べてみます。

データに含まれている1つ1つの数値から平均値をひいた値を、それぞれの数値の といいます。

A組、B組のそれぞれについて、次の表の空欄を埋めてください。

<A組>

点数	0	1	2	3	4	5	6	7	8	9	10
偏差											
人数	1	0	0	0	2	4	2	0	0	0	1
偏差×人数											

<B組>

点数	0	1	2	3	4	5	6	7	8	9	10
偏差											
人数	2	1	0	0	0	4	0	0	0	1	2
偏差×人数											

したがって、「偏差×人数」の合計は

A組…… , B組……

となります。

◎ 以上からわかること

(5) (4)の問題点を克服するために、偏差の絶対値を考えてみます。次の表の空欄を埋めてください。

<A組>

点数	0	1	2	3	4	5	6	7	8	9	10
偏差の絶対値											
人数	1	0	0	0	2	4	2	0	0	0	1
偏差の絶対値×人数											

<B組>

点数	0	1	2	3	4	5	6	7	8	9	10
偏差の絶対値											
人数	2	1	0	0	0	4	0	0	0	1	2
偏差の絶対値×人数											

したがって、「偏差の絶対値×人数」の合計は

A組…… , B組……

となって、これならA組とB組の区別ができそうです。

(6) 「偏差の2乗」を考えることもできます。次の表の空欄を埋めてください。

<A組>

点数	0	1	2	3	4	5	6	7	8	9	10
偏差の2乗											
人数	1	0	0	0	2	4	2	0	0	0	1
偏差の2乗×人数											

<B組>

点数	0	1	2	3	4	5	6	7	8	9	10
偏差の2乗											
人数	2	1	0	0	0	4	0	0	0	1	2
偏差の2乗×人数											

したがって、「偏差の2乗×人数」の合計は

A組…… , B組……

となります。

ここで、上で求めた「偏差の2乗×人数」の合計を、人数でわってみます。つまり、「偏差の2乗の平均」を求めるということです。

A組…… , B組……

いま求めた、「偏差の2乗の平均」を といいます。

◎ どんな統計量？

(7) A組, B組のそれぞれについて, (6)で計算した統計量の**平方根**をとってみましょう。

2乗して x になる数のことを, x の平方根と呼びます。例えば, 16 の平方根は 4 と -4 の2つです。一般に, 正の数の平方根は 0 を除き, 正と負の2つずつあります (負の数の平方根は考えません)。

x の平方根のうち, 正であるものを \sqrt{x} , 負であるものを $-\sqrt{x}$ で表します。上の例で言えば

$$\sqrt{16}=4, \quad -\sqrt{16}=-4$$

です。16 のように, 整数の2乗となっている数 (**平方数**という) は $\sqrt{\quad}$ の記号を使わずに表すことができますが, そうでない場合は, $\sqrt{\quad}$ の記号を残したままでかまいません。例えば, 5 の平方根は

$$\sqrt{5} \text{ と } -\sqrt{5}$$

です。 $\sqrt{\quad}$ の記号を使わないと表せない数は, それに近い値として小数で表すことができます。例えば, $\sqrt{5}$ は約 2.236 です。

では, A組, B組のそれぞれの正の平方根を, 電卓などを使って求めてみてください。

A組…… , B組……

このように, 分散の平方根のうち, 負でない方を といいます。

◎ どんな統計量？

これまで学習してきたことを、文字を使ってまとめておきましょう。

n 個の数値があるとします。この n 個の数値のまとまりを**データ**と呼びます。データに含まれている n 個の数値を

$$x_1, x_2, \dots, x_n$$

と表すことにします。

◎ 平均値

$$\mu = \frac{x_1 + x_2 + \dots + x_n}{n}$$

μ はギリシャ文字で「**ミュー**」と読みます。

を、このデータの**平均値**と呼びます。

◎ 分散

n 個のデータ x_1, x_2, \dots, x_n の平均値を μ で表すこととします。このとき、数値 x_1 と平均 μ の差、つまり

$$x_1 - \mu$$

を、 x_1 の**偏差**と呼びます。 x_1 と同様にして、 x_2, x_3, \dots, x_n の偏差をそれぞれ考えることができます。

x_1, x_2, \dots, x_n の偏差をそれぞれ 2 乗した値の平均値を、**分散**と呼びます。すなわち、分散を V とすると

$$V = \frac{(x_1 - \mu)^2 + (x_2 - \mu)^2 + \dots + (x_n - \mu)^2}{n}$$

です。

◎ 標準偏差

分散の平方根のうち、負でない方を**標準偏差**といいます。すなわち、分散を V とし、標準偏差を σ とすると

$$\sigma = \sqrt{V}$$

σ はギリシャ文字で「**シグマ**」と読みます。 Σ の小文字です。

です。

では、これらを踏まえて計算の練習をしてみましょう。

問題 3

5 人の小テストの得点のデータが、次のように与えられています。

$$5, 7, 5, 10, 8$$

このデータの分散 V と標準偏差 σ を求めてください。

第 1 章

第 2 章

第 3 章

第 4 章

第 5 章

分散が「平均値からの隔たりの度合い」つまり「データのバラツキのしかた」を表すのはすでに学習した通りですが、ここで疑問が生じます。

◎ 結局、分散を知ることによってどのような意味があるのだろうか？

◎ 分散を知ると何が嬉しいのだろうか？

◎ バラツキの度合いが大きい小さいかを論じることによってどのような意義があるのだろうか？

ということです。

これらのことを考えるために、次の問題を考えてみましょう。

問題4

太郎くんは、100点満点のテストを受けました。その結果

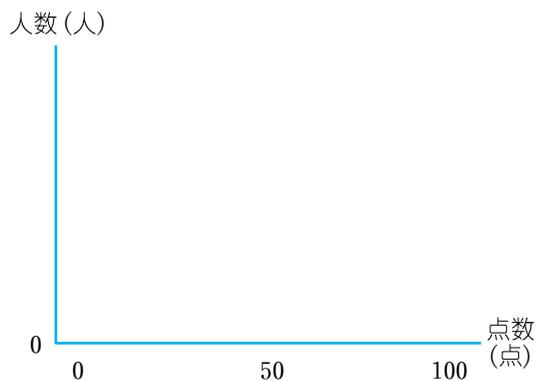
クラスの平均点が50点で、自分の点数が100点である

ことがわかりました。

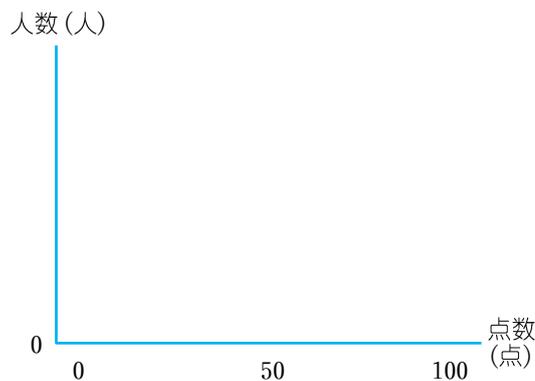
太郎くんは喜んでいきます。

……が、これは本当に喜んでいいことなのでしょうか？

<ケース1>



<ケース2>



◎ 分散を考えることの意義

★おまけコーナー★

分散について、2つ補足しておきます。計算がやや難しいので、数学好きな人はどうぞ。

(1) 分散の計算をラクに

分散を定義通りに計算するとちょっと大変です。そこで、もう少し簡単に計算できる方法はないか考えてみます。

話を簡単にするために、2つのデータ a, b で分散を考えます。この2つのデータの平均値 μ は

$$\mu = \frac{a+b}{2}$$

だから、分散 V を定義通りに計算すると

$$V = \frac{1}{2}\{(a-\mu)^2 + (b-\mu)^2\}$$

となります。少し難しいのですが、この右辺を変形していきましょう。

$$\begin{aligned} V &= \frac{1}{2}\{(a^2 - 2a\mu + \mu^2) + (b^2 - 2b\mu + \mu^2)\} \\ &= \frac{1}{2}(a^2 + b^2) - \frac{1}{2} \times 2\mu(a+b) + \frac{1}{2} \times 2\mu^2 \\ &= \frac{a^2 + b^2}{2} - 2\mu \times \mu + \mu^2 \\ &= \frac{a^2 + b^2}{2} - \mu^2 \end{aligned}$$

この結果により、分散は

《2乗の平均》-《平均の2乗》

で計算できることがわかります。 **問題3** をこの公式を使って解いてみましょう。

(2) なぜ偏差の2乗を使うのか

p.21 の(5)で考えたように、分散を「偏差の2乗の平均」で定義せず、「偏差の絶対値の平均」で定義することはできないのでしょうか。

例として、3つのデータ $(x_1, x_2, x_3) = (1, 2, 6)$ について考えてみます。これらのデータの平均値 μ は

$$\mu = \frac{x_1 + x_2 + x_3}{3} = \frac{1 + 2 + 6}{3} = 3$$

なので、「偏差の絶対値の平均」 V' を計算すると

$$\begin{aligned} V' &= \frac{1}{3}(|x_1 - \mu| + |x_2 - \mu| + |x_3 - \mu|) \\ &= \frac{1}{3}(|1 - 3| + |2 - 3| + |6 - 3|) \\ &= \frac{1}{3}(2 + 1 + 3) = 2 \end{aligned}$$

となります。

ここで、 μ の代わりに3つのデータの中央値2を使って上の V' を計算してみます。

$$V' = \frac{1}{3}(|1 - 2| + |2 - 2| + |6 - 2|) = \frac{5}{3}$$

となり、上で計算したものよりも小さくなります。

一般に、 x の関数として

$$\frac{1}{n}(|x_1 - x| + |x_2 - x| + \dots + |x_n - x|)$$

を考えると、この関数は x が x_1, x_2, \dots, x_n の中央値に一致するとき最小値をとることが知られています。分散の意味からすると、 x が x_1, x_2, \dots, x_n の平均値に一致するとき最小値をとってほしいわけですが、そうはならないので、分散の定義として「偏差の絶対値の平均」が使われないのです。ちなみに、「偏差の絶対値の平均」は**平均偏差**と呼ばれています。

第1章
第2章
第3章
第4章
第5章

3

Excel を使ってみよう

表計算ソフトを利用して、効率よくデータを分析しましょう。

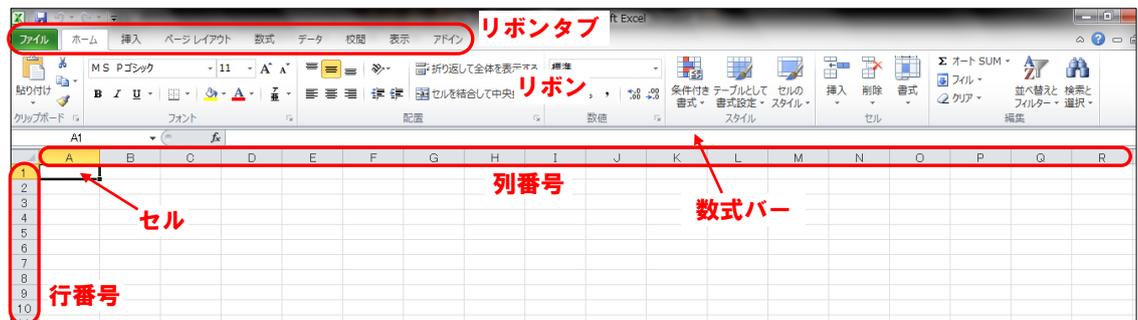
これまで、いろいろな統計量を計算してきましたが、手計算でそれを行うのは少々面倒です。そこで、大量のデータを扱うのに便利な PC のアプリケーションを紹介します。それが Excel (エクセル) です。

Excel は、Windows でなじみの米国 Microsoft 社が開発したアプリケーションです、**表計算ソフト**と呼ばれるものの 1 つに分類されます。表計算ソフトとは、データを入力して様々な方法で集計したり、数式を入力して簡単に計算したりすることができるソフトです。

世の中にはいろいろな表計算ソフトがありますが、Excel は計算機能に優れ、複雑な計算も瞬時にできる上、データをもとにグラフをつくることもできます。

1. 画面の説明

Excel を立ち上げると、次のような画面になります。各部の名称は次の通りです。



① リボントブ

「ファイル」、「ホーム」、「挿入」、「ページレイアウト」、「数式」、「データ」、「校閲」、「表示」、「アドイン」などを切り替えられます。

② リボン

いろいろな操作をするためのボタンや、ドロップダウンリストが機能ごとに各タブに格納されています。

③ セル

データを入力できる場所で、文字や数字、数式などを入力できます。

④ 数式バー

選択 (クリック) したセルの内容が表示されます。数式バーをクリックすると内容を編集できます。

⑤ 行番号・列番号

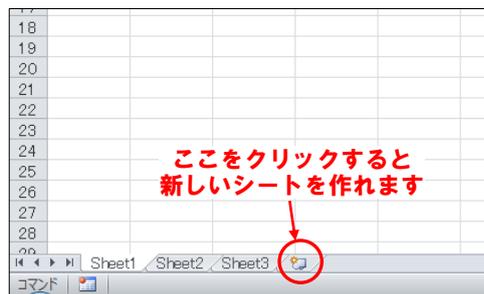
行番号は数字で行の位置を、列番号はアルファベットで列の位置を示します。列番号は、A から始まり Z の次は AA, AB, ……のように続きます。

各セルには、番地が割り振られており、行番号・列番号を組み合わせ、「A1」、「B6」などと表します (例えば、上の図の**アクティブセル** (太枠で囲まれたセル) は A1 セルです)。

2. ブックとワークシート

Excel の表のファイル全体のことを**ブック**といいます。Excel を立ち上げると、画面の左下に右図のように「Sheet 1」, 「Sheet 2」, ……などと表示されていることがわかります。

これらを**ワークシート**と呼びます。それぞれのワークシートで別々の作業をすることができるし、ワークシート間で関連するデータを扱うこともできます。



3. データの入力

(1) 入力の基本

キーボードを押して文字を入力したとき、アクティブセルに入力されます。

B 列に文字を入力したのが下の図です。「太郎」, 「次郎」, 「三郎」についてはセルの枠内に収まっていますが、「レオナルド・ダ・ヴィンチ」は枠からはみ出してしまっています。これを枠内に収めたいときは、列番号の右端をダブルクリックします。



(2) 入力の修正

入力した文字を修正するには、次の 3 通りの方法があります。

- ① 修正したいセルをダブルクリックする。
- ② 修正したいセルをクリックした後、キーボードの **F2** を押す。
- ③ 修正したいセルをクリックした後、数式バーをクリックする。

セルに入力した文字などを削除したい場合は、キーボードの **Delete** キーを押せば OK です。

(3) 連続したデータの入力

Excel には入力を簡略化するため、**オートフィル**という機能が備わっています。

例えば、右の図で A 列の「番号」の下に「太郎」から「松子」まで通し番号を振りたい場合、1 から 10 までの数を直接入力するのは大変なので、オートフィルの機能を使って 10 までの数を簡単に入力することができます。

アクティブセルの太枠の右下の角 (**フィルハンドル**と呼びます) にマウスカーソルをおくと、図のような十字のマークに変わります。

このマークが変わったら、角をクリックしたまま値を入力したい方向にマウスを動かすと、右の図のようになり、1 から 10 までの数字があっという間に入力できます。

Excel でデータを処理するとき、高頻度で使う機能なので、覚えておきましょう。



第 1 章

第 2 章

第 3 章

第 4 章

第 5 章

4. データの処理～数式の使用～

(1) 計算と数式

Excel が「表計算ソフト」と呼ばれるように、その中心となるのは「計算機能」です。Excel で計算・数式を扱うには、以下のような決まりがあります。

① 数式（計算式）の入力は、**数式バー**で行う。

② 数式は、必ず「=」（半角のイコール）から始め、「**=100+200+300**」のように入力します（ちなみに、このように入力したあと **Enter** キーを押すと、そのセルには「**600**」と表示されます）。

③ 通常は、数式バーには入力した数式が表示され、その答えがセルに表示されます。

※ この例では、加法だけですが、減法の場合は「**=200-100**」、乗法の場合は「**=200*100**」、除法の場合は「**=200/100**」と入力します。

(2) 得点の合計を計算してみよう

右の図の G3 セルから G5 セルにかけて、太郎、次郎、三郎の 3 教科の得点の合計を入力したい場合、数式を利用することができます。

① 直接入力の場合

G3 セルをクリックして、数式バーに「**=55+23+86**」と入力すると、計算結果 **164** が G3 セルに表示されます。

もっとも単純な方法ですが、例えば、あとで得点の入カミスが発覚して一部のデータを修正しなければならない事態になったとき、合計までは修正されないの、もう一度同じように入力しなければならず面倒です。

番号	氏名	数学の得点	英語の得点	社会の得点	合計
1	太郎	55	23	86	164
2	次郎	98	46	21	
3	三郎	12	68	89	

② セルの参照機能を利用する場合

右の図のように、G3 セルをクリックして、数式バーに「**=D3+E3+F3**」と入力しても①と同じ結果が得られます。これは、『D3, E3, F3 に入力されている値をすべて足しなさい』という命令になるので、あとで得点を修正するようなことがあっても、修正後に合計の値も追従して変更されます。

番号	氏名	数学の得点	英語の得点	社会の得点	合計
1	太郎	55	23	86	164
2	次郎	98	46	21	
3	三郎	12	68	89	

③ 関数を使う場合

Excel には**関数**と呼ばれる強力な計算機能が用意されています。関数とは、目的の計算をより簡単にできるようあらかじめ Excel に用意されている数式のことです。これを使えば、面倒な計算も簡単に行うことができます。

例えば、複数の値の合計を計算する「**SUM 関数**」という関数があり、上の図のように G3 セルに「**=SUM(D3:F3)**」

と入力すると、①、②と同じ結果が得られます。「**D3:F3**」は『D3 セルから F3 セルまでの範囲』を意味します。これは

「**=SUM(D3,E3,F3)**」

と入力しても同じです（ただし、この方法はセルが多くなると大変です）。

番号	氏名	数学の得点	英語の得点	社会の得点	合計
1	太郎	55	23	86	164
2	次郎	98	46	21	
3	三郎	12	68	89	

5. 関数を使ってみよう

(1) 関数の種類

リボンの「数式」タブの「関数の挿入」をクリックすると、関数の一覧を見ることができます。関数は扱うデータの種類に応じて、いくつかに分類されています（例えば、数値を扱う関数は「数学/三角」という分類に、文字列を扱う関数は「文字列操作」という分類に入っています）。

第1章

第2章

第3章

第4章

第5章

(2) 関数の基本構造

さきほどのSUM関数を例に説明します。

SUM(**)**
引数

SUM関数の場合、かっこ内にはセル番地（セル範囲）を入力します。

関数は、一般的にその関数名のあとにかっこをつけて、かっこ内にその関数が必要とするものを入力します。かっこ内に入れるべき値は**引数**と呼ばれ、その種類は数値、文字列、セル参照などさまざまです。

(3) 関数を使ってみよう

セルに関数を入力する場合、2通りの方法があります。

① 関数のリストから選ぶ方法

リボンの「数式」タブから「関数の挿入」を選択すると関数の一覧が表示されるので、その中から必要な関数を選んでセルに挿入することができます。

例えば、さきほどのSUM関数は、「数学/三角」の分類にあるので、その中から「SUM」を選択し、右下のOKをクリックすると、下の図の右側のようなウィンドウが表示されます。



その関数が必要としている引数（この場合はセル範囲）を入力し、右下のOKをクリックすると、シート上の選択したセルに関数が挿入されます。

② 直接入力する方法

数式バーに関数を直接入力することもできます。慣れてくるとこちらの方が圧倒的に早く処理できます。

(4) よく使う関数の例

SUM関数も非常によく使いますが、それ以外によく使う関数を挙げておきます。

① AVERAGE関数

指定した範囲の数値の平均値を求める関数です。右の図のD6セルに、3人の数学の得点の平均点を表示したい場合

[=AVERAGE(D3:D5)]

と入力すればOKです。なお、指定した範囲に空白があった場合は無視されます。

	A	B	C	D	E	F	G	H
1								
2		番号	氏名	数学の得点	英語の得点	社会の得点	合計	
3		1	太郎	55	23	86	164	
4		2	次郎	98	46	21	165	
5		3	三郎	12	68	89	169	
6			平均点	55				
7								

② RANK 関数

指定した数値が、指定した範囲の中で第何位になるのかを求める関数です。

右の図の H3 セルに、太郎の合計点が 3 人のうち何位になるのかを表示したい場合

		H3							
		=RANK(G3,G3:G5,0)							
	A	B	C	D	E	F	G	H	I
1									
2		番号	氏名	数学の得点	英語の得点	社会の得点	合計	順位	
3	1	太郎	55	23	86	164	3		
4	2	次郎	98	46	21	165			
5	3	三郎	12	68	89	169			
6		平均点	55						
7									

「=RANK(G3,G3:G5,0)」

と入力します。引数がいくつかある場合、上のように「,」で区切ります。この場合

『G3 セルの値が、G3 から G5 までの範囲で第何位かを表示しなさい』

という意味になります。3 番目の引数「0」は、順位を得点の高い方を 1 位として第何位になるかを表示するよう指定するためのものです（これを「1」にすると、得点の低い方を 1 位とした場合が表示されます）。なお、数値が同じ場合は同率の順位がつきます。

③ MIN 関数, MAX 関数

指定した範囲の最小値、最大値を求める関数です。

右の図の D7 セルには

「=MAX(D3:D5)」

と入力されています。

		D6							
		=MAX(D3:D5)							
	A	B	C	D	E	F	G	H	I
1									
2		番号	氏名	数学の得点	英語の得点	社会の得点	合計	順位	
3	1	太郎	55	23	86	164			
4	2	次郎	98	46	21	165			
5	3	三郎	12	68	89	169			
6		最高点	98						
7		最低点	12						
8									

④ COUNTIF 関数

指定した範囲に、指定した値（文字列）がいくつあるかを求める関数です。

例えば、右の図の E12 セルに、男子が何人いるか表示したい場合

「=COUNTIF(E4:E11,"男")」

と入力します（上のように文字列の個数を数える場合、ダブルクォーテーションで囲む必要があります）。

		E12					
		=COUNTIF(E4:E11,"男")					
	A	B	C	D	E	F	
1							
2							
3					性別		
4				太郎	男		
5				花子	女		
6				松子	女		
7				次郎	男		
8				三郎	男		
9				五郎	男		
10				桃子	女		
11				梅子	女		
12				男性	4	人	
13				女性		人	
14							

⑤ MEDIAN 関数, MODE 関数

中央値および最頻値を求める関数です。いずれも引数として、**MEDIAN(セル範囲)**、**MODE(セル範囲)**として使います。

⑥ VAR 関数, STDEVP 関数

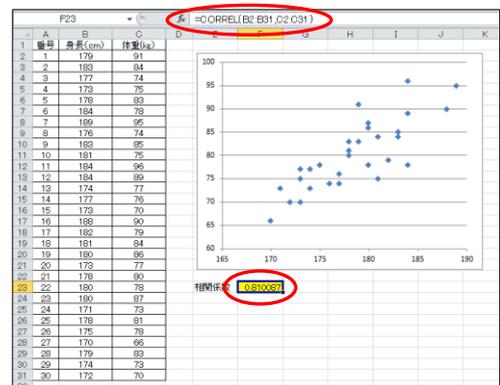
分散、標準偏差を求める関数です。いずれも引数として、**VAR(セル範囲)**、**STDEVP(セル範囲)**として使います。

⑦ CORREL 関数

相関係数を求める関数です。引数は 2 つで、第 1 引数と第 2 引数のそれぞれに散布図の横軸、縦軸に使用しているデータ群が入力されているセル範囲を入力します。右の図の F23 セルには

「=CORREL(B2:B31,C2:C31)」

と入力されています。



ここからは、少しレベルが高くなりますが、よく使うものなので今のうちに覚えておきましょう。

⑧ IF 関数

指定した条件（論理式）によって、対象が真（TRUE）か偽（FALSE）かを判定し、それぞれに指定した値を求める関数です。

書式は、**IF(論理式, 真の場合, 偽の場合)**となります。言葉だけではわからないと思うので、実際の使い方を見てみましょう。

右の図で、E 列に、数学の得点が 60 点以上なら「合格」、60 点未満なら「不合格」と表示させたいとします。このとき

「=IF(D3>=60,"合格","不合格")」

ように入力すれば OK です。つまり、「**D3>=60**」（D3 は 60 以上）という式が真ならば、「合格」と表示させ、偽ならば「不合格」と表示させる指示になっています。太郎の数学の得点は 55 点で、「**D3>=60**」を満たさない（つまり偽である）ので、第 3 引数の「不合格」が適用され、「不合格」が表示されています。

ちなみに、IF 関数は、関数の構文の中にさらに IF 関数を入れることができます（こういう構造を、**入れ子構造**といいます）。

例えば、右の図の数式バーを見ると、IF 関数の中に IF 関数を入れていますが、これは、『数学が 60 点以上の場合は（国語が何点であっても）「合格」、数学は 60 点未満だが国語は 60 点以上なら「数学がんばれ」、両方 60 点未満なら「不合格」と表示せよ』という命令になります。

⑨ VLOOKUP 関数

リストや表の指定した範囲の中で、左端の列内から指定した文字や数を検索し、指定した列からそれに該当するデータを抽出する関数です。

これも例を見てみましょう。

右の図で、H2 セルに名前を入力すると、I2 セルにその人の数学の得点が表示されるようにしたいとき

「=VLOOKUP(H2,C3:E5,2,FALSE)」

と入力します。この意味は

『**H2** に入力された名前を、範囲 **C3:E5** の左端の列 C から探し、その範囲の **2** 列目の文字（数値）を求める』

となります（「FALSE」は、H2 に入力したものと完全一致したものを検索することを意味しています。ここを「TRUE」にした場合は完全一致ではなく部分一致となります）。

	A	B	C	D	E	F	G
1							
2		番号	氏名	数学の得点	合格		
3		1	太郎	55	不合格		
4		2	次郎	98	合格		
5		3	三郎	12	不合格		
6							

	A	B	C	D	E	F	G	H
1								
2		番号	氏名	数学の得点	国語の得点	合格		
3		1	太郎	55	65	数学がんばれ		
4		2	次郎	98	24	合格		
5		3	三郎	12	30	不合格		
6								

	A	B	C	D	E	F	G	H	I	J
1									得点	
2		番号	氏名	数学の得点	国語の得点		氏名	太郎	55	
3		1	太郎	55	65					
4		2	次郎	98	24					
5		3	三郎	12	30					
6										

第 1 章

第 2 章

第 3 章

第 4 章

第 5 章

6. 相対参照と絶対参照

少し高度な話になりますが、これを知っているか知らないかで作業効率が大きく変わっていきますので、ぜひ知っておきたい内容です。

Excel のセル参照は、大きく分けて、「絶対参照」と「相対参照」の 2 通りの方法が用意されています。

① 絶対参照……参照するセル番地が常に固定される参照方式

→「\$A\$1」のように「\$」をつけることによって絶対参照となり、数式をコピーすると、どの数式のおいても同一のセルを参照します。

② 相対参照……参照先が数式に連動して変化する参照方式

→数式をコピーすると、コピー先のセル位置に応じて参照先のセルが自動的に変化します。

具体例で確認しましょう。

<相対参照>

1. D4 に「=B4 * C4」を入力します。
2. フィルハンドルを、D6 までドラッグします。

	A	B	C	D	E	F	G	H
1							値引率	20%
2								
3	商品名	単価	販売数	金額	値引き	差引金額		
4	商品A	4000	3	12000				
5	商品B	5000	2					
6	商品C	2800	4					

ここで、D5、D6 セルに表示されている数式を見ると、参照先が自動的に変化していることがわかります。

	A	B	C	D	E	F	G	H
1							値引率	20%
2								
3	商品名	単価	販売数	金額	値引き	差引金額		
4	商品A	4000	3	12000	=B4 * C4			
5	商品B	5000	2	10000	=B5 * C5			
6	商品C	2800	4	11200	=B6 * C6			

<絶対参照>

1. E4 に「=D4 * H1」を入力します。

	A	B	C	D	E	F	G	H
1							値引率	20%
2								
3	商品名	単価	販売数	金額	値引き	差引金額		
4	商品A	4000	3	12000	2400			
5	商品B	5000	2	10000				
6	商品C	2800	4	11200				

2. フィルハンドルを、E6 までドラッグします。正常な計算結果が出ずに、「商品 B」、「商品 C」は「0」となっていました。
3. 計算式を確認します。

	A	B	C	D	E	F	G	H
1							値引率	20%
2								
3	商品名	単価	販売数	金額	値引き	差引金額		
4	商品A	4000	3	12000	2400	=D4 * H1		
5	商品B	5000	2	10000	0	=D5 * H2		
6	商品C	2800	4	11200	0	=D6 * H3		

現在、E4 から E6 に入っている数式は、

$$=D4 * H1$$

$$=D5 * H2$$

$$=D6 * H3$$

となっています。これは、もともと E4 に入っていた数式を下にコピーしたので、さっきの図の相対参照が働き、セル参照が1つずつ下にずれているのです。

ここでは、「値引率」のセルは H1 だけなので、1つずつずれることなく、

$$=D4 * H1$$

$$=D5 * H1$$

$$=D6 * H1$$

のように常に H1 を参照するように設定します。

4. あらためて、E4に「**=D4*\$H\$1**」と入力します。これで参照先の H1 が固定されます。
5. この状態でフィルハンドルを E6 までドラッグすると、金額に応じた値引額が正常に表示されます。

	A	B	C	D	E	F	G	H
1							値引率	20%
2								
3	商品名	単価	販売数	金額	値引き	差引金額		
4	商品A	4000	3	12000	2400	数式「=D4*\$H\$1」		
5	商品B	5000	2	10000	2000	数式「=D5*\$H\$1」		
6	商品C	2800	4	11200	2240	数式「=D6*\$H\$1」		
7								

第1章

第2章

第3章

第4章

第5章

このように、数式をコピーするときは参照先を連動させる必要があるのかないのかを考えておかなければなりません。

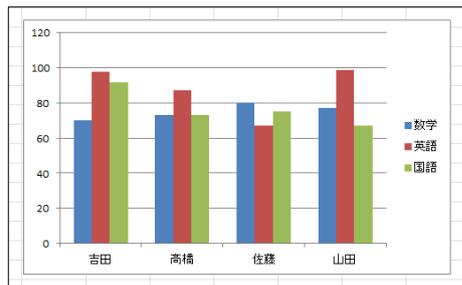
7. グラフをつくってみよう

Excel には、つくったデータを元にして、棒グラフや円グラフなど、いろいろなグラフをつくる機能が備わっています。

ここでは、

	A	B	C	D	E	F	G
1							
2			数学	英語	国語	合計	
3		吉田	70	98	92	260	
4		高橋	73	87	73	233	
5		佐藤	80	67	75	222	
6		山田	77	99	67	243	
7							

という表を元にして



というグラフをつくる手順を説明します。

① 表のグラフ化したい部分を選択する

ここでは、各生徒の数学・英語・国語の得点を棒グラフにしたいので、項目も含まれるように B2 から E6 までの範囲を長方形になるように選択します。



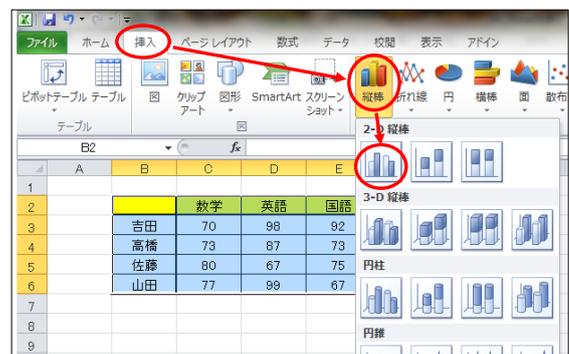
② リボンの「挿入」タブから表示したいグラフを選択する

グラフに必要な範囲を選択したら、リボンの「挿入」タブから、「縦棒」を選びます。

そうすると、棒グラフの種類が選択できるようになるので、ここでは「2-D 縦棒」のいちばん左のものを選びましょう。

……実は手順はこれだけです。たった 2 ステップで、さきほどのグラフをつくることができます。

あとは、このグラフのレイアウトなどに手を加えたい場合に次のステップに進みます。



③ レイアウトを変える



グラフの軸を変更したり、棒の色を変えたり、パーツをつけ足したりしたい場合は、できあがったグラフを選択した状態で、リボンタブとして加わった「グラフツール」を使って修正していきます。

好みに合わせて、いろいろ試してみてください。

さらに学習したい人は、こちらを試してみましょう。

8. 分析ツールを使ってみよう

せっかく Excel を使うので、1年生のときに学習した度数分布表やヒストグラムも作れたらいいですよ。そこで、ここでは発展編として「分析ツール」を紹介します。

ただ、この機能は Excel の初期状態で装備されているものではないので、少し準備が必要です。

まず、リボンタブの「データ」をクリックしてください。



上の図のように、「データ」タブの右端に「ソルバー」、「データ分析」と表示されていれば準備完了です。もしこの表示がなければ、以下の作業をしてください。

- ① リボンの「ファイル」タブをクリックし、「オプション」を選択
- ② 左端のメニューから「アドイン」を選択
- ③ 下方の「設定」をクリック
- ④ 「有効なアドイン」の一覧から、「ソルバーアドイン」と「分析ツール」にチェックを入れ、OK をクリック

これで分析ツールが有効になります。

9. ヒストグラムをつくってみよう

ここでは、あるクラスの生徒 35 人の 50 m 走の記録について、1 で導入した分析ツールを使って、階級の幅 0.5 秒のヒストグラムをつくってみます。

(1) データの準備

右の図のように、生徒それぞれの 50 m 走の記録を示した表を用意します。

① 階級

ここでは、階級の幅を 0.5 秒にし

「6.5 秒以上 7.0 秒未満」

「7.0 秒以上 7.5 秒未満」

「7.5 秒以上 8.0 秒未満」

「8.0 秒以上 8.5 秒未満」

「8.5 秒以上 9.0 秒未満」

「9.0 秒以上 9.5 秒未満」

の 6 つの階級をつくることを考えます。

この後の操作の都合上、D 列に(階級の最小値-0.1)から(階級の最大値-0.1)までの数値を入力しておきます。

Excel の仕様上、例えば 6.5 秒のような階級の境界にある値は、「6.5 秒以上 7.0 秒未満」の階級ではなく、「6.0 秒以上 6.5 秒未満」の階級に入ってしまう。上で 0.1 をひいているのは、6.5 秒が「6.5 秒以上 7.0 秒未満」の階級に入るようにするための措置です。

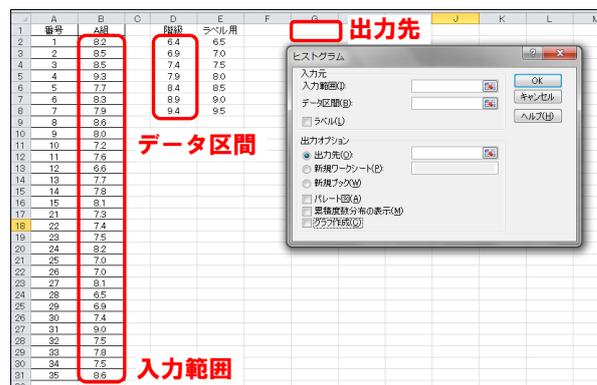
② ラベル用

このままだとヒストグラムをつくったときに横軸に表示される値が D 列に入力したものになってしまうので、横軸のラベルを変更するために、E 列に表示させたい実際の数値を入力しておきます。

	A	B	C	D	E	F
1	番号	A組		階級	ラベル用	
2	1	8.2		6.4	6.5	
3	2	8.5		6.9	7.0	
4	3	8.5		7.4	7.5	
5	4	9.3		7.9	8.0	
6	5	7.7		8.4	8.5	
7	6	8.3		8.9	9.0	
8	7	7.9		9.4	9.5	
9	8	8.6				
10	9	8.0				
11	10	7.2				
12	11	7.6				
13	12	6.6				
14	13	7.7				
15	14	7.9				
16	15	8.1				
17	21	7.3				
18	22	7.4				
19	23	7.5				
20	24	8.2				
21	25	7.0				
22	26	7.0				
23	27	8.1				
24	28	6.5				
25	29	6.9				
26	30	7.4				
27	31	9.0				
28	32	7.5				
29	33	7.8				
30	34	7.5				
31	35	8.6				
32						

(2) ヒストグラムをつくってみよう

ここまで準備ができれば、リボンの「データ」タブの右端に表示されている「データ分析」をクリックします。表示されるメニューから「ヒストグラム」を選択し、OKをクリックすると、右の図のようなウィンドウが出ます。



① 入力範囲

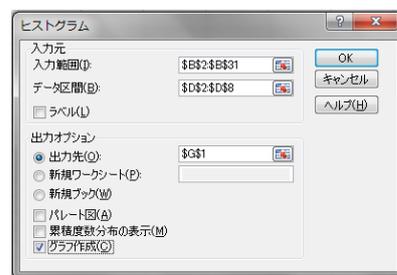
実際にグラフ化したいデータの数値を選択します。をクリックして、B2 から B31 までの範囲を選択してください。

② データ区間

階級を決めるためのデータの数値を選択します。をクリックして D2 から D8 までの範囲を選択してください。

③ 出力先

度数分布表を出力するセルを選択します。ここでは、G1 セルを指定します。

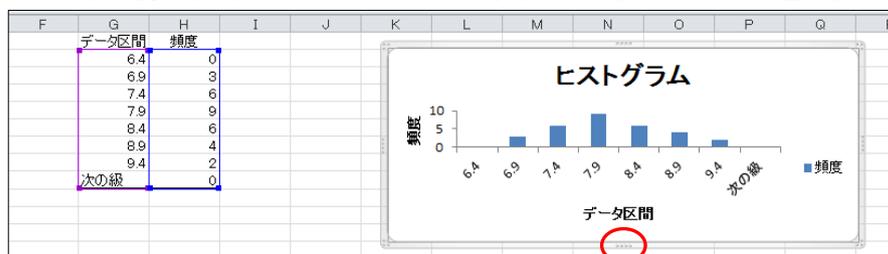


④ グラフ作成

チェックボックスにチェックを入れます。

ここまでの作業が完了すると、右の図のようになります。

OK をクリックすれば、度数分布表とヒストグラムが表示されます。それが下の図です。



このままだとヒストグラムが異様に小さくなっているので、上の赤丸部分をドラッグし、グラフの大きさを変えます。

(3) ヒストグラムのレイアウトを修正しよう

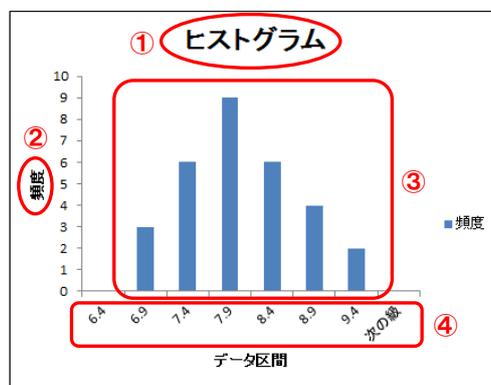
変更したいところは、右の図の①～④の部分です。

① タイトル

タイトル部分をクリックすると編集できます。例えば、「50 m 走の記録」などとするとよいでしょう。

② 縦軸の名称

クリックすると編集できます。ここでは、「人数」としておきます。



③ グラフ

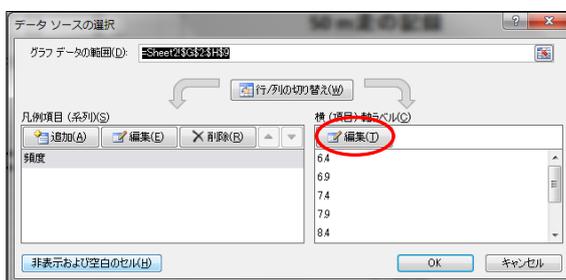
本来ヒストグラムは棒と棒の間隔がないので、棒の間隔を 0 にします。グラフ部分をクリックして、リボンに表示される「デザイン」タブの「グラフのレイアウト」の中から「レイアウト 8」を選択してください。



④ 横軸のラベル

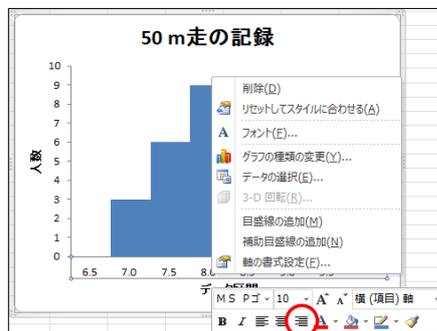
このままだと D 列に入力した階級が表示されているので、ここを E 列に入力した数値に変えます。横軸のラベル上で右クリックし、「データの選択」を選ぶと、右のようなウィンドウが出ます。

「編集」をクリックし、E2 から E8 までの範囲を選択すると、データラベルが変更されます。



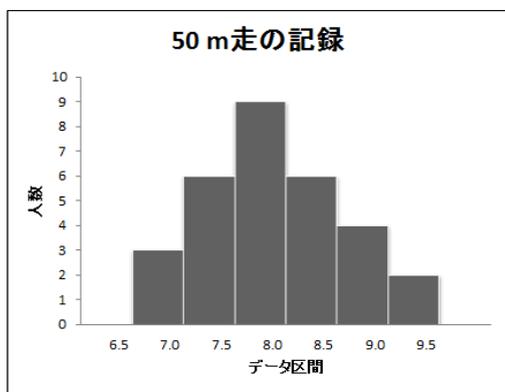
このままだとラベルの位置が棒の真ん中になっていて、通常のヒストグラムのようにないないので、ラベルの表示位置を変更します。

ラベルの上で右クリックすると右のようなメニューが出るので、赤丸をつけた部分をクリックしてください。



⑤ あとはデザインをいろいろ変えてみる

右クリックで表示されるメニューから「〇〇の書式設定」を選択すると、デザインをいろいろ変更することができます。あとは好みでいろいろいじってみてください。で、ひとまず完成したのが下の図です。



練習問題

ここまで学習してきたことを、下の問題を通じて練習してみましょう。Teams にアップロードされている Excel ファイルをダウンロードし、自身のタブレットでチャレンジしてみてください。

問題ごとにシートが分かれているので、チャレンジしたい問題のシートを選択して取り組んでください。

問題 1 (Sheet 1)

D 列の黄色で塗ったセルに、各日付の「残席」の欄に、定員から各日付の申込者数を引いた数が表示されるようにしてください。

	A	B	C	D	E
1	日別申込者数一覧				
2			定員	40	
3					
4	日付	曜日	申込者数	残席	
5	8月20日	月	16		
6	8月21日	火	37		
7	8月22日	水	9		
8	8月23日	木	40		
9	8月24日	金	25		
10					

問題 2 (Sheet 2)

D, E 列と 5, 6 行目の各合計と平均を、関数を使って表示してください。

	A	B	C	D	E
1	商品	上半期	下半期	合計	平均
2	りんご	5800	2790		
3	みかん	4580	4680		
4	バナナ	7540	7720		
5	合計				
6	平均				
7					

問題 3 (Sheet 3)

(1) E2, E3 および E5, E6 セルに、男子（番号 1~15）、女子（番号 21~35）の平均値、中央値が表示されるように関数を入力してください。

(2) 適切な関数を入力し、E8 セルに、A 組の中で 50 m 走の記録が 8.0 秒以下の人数を表示させてください。

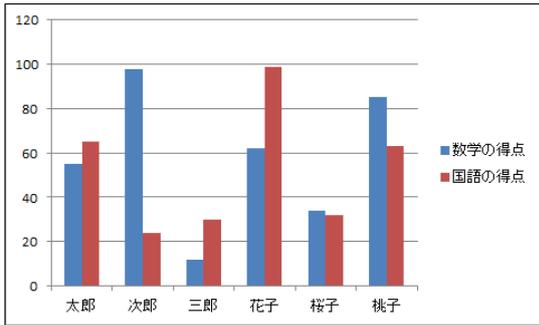
	A	B	C	D	E
1	番号	A組			
2	1	8.2	男子平均		
3	2	8.5	男子中央値		
4	3	8.5			
5	4	9.3	女子平均		
6	5	7.7	女子中央値		
7	6	8.3			
8	7	7.9	8.0秒以下の人数		
9	8	8.6			
10	9	8.0			
11	10	7.2			
12	11	7.6			
13	12	6.6			
14	13	7.7			
15	14	7.8			
16	15	8.1			
17	21	7.3			
18	22	7.4			
19	23	7.5			
20	24	8.2			
21	25	7.0			
22	26	7.0			
23	27	8.1			
24	28	6.5			
25	29	6.9			
26	30	7.4			
27	31	9.0			
28	32	7.5			
29	33	7.8			
30	34	7.5			
31	35	8.6			
32					

問題4 (Sheet 4)

(1) G1セルに「男子」か「女子」を入力し、G2セルに名前を入力すると、H2セルにその人の数学の得点が表示されるように関数を入力してください。

(2) この表を利用して、下のようなグラフをつくってください。

	A	B	C	D	E	F	G	H
1	番号	氏名	数学の得点	国語の得点		性別	女子	
2	1	太郎	55	65		氏名	桜子	
3	2	次郎	98	24				
4	3	三郎	12	30				
5	1	花子	62	99				
6	2	桜子	34	32				
7	3	桃子	85	63				
8								



- 第1章
- 第2章
- 第3章
- 第4章
- 第5章

問題5 (Sheet 5)

右の表は、あるクラスの生徒 30 人の身長と体重のデータをまとめたものです。

この表を元にして、身長を横軸、体重を縦軸として散布図を作成し、適切な関数を使って F2セルに相関係数を表示してください。

	A	B	C	D	E	F	G
1	番号	身長(cm)	体重(kg)				
2	1	179	91		相関係数		
3	2	183	84				
4	3	177	74		↓ここに散布図を表示させる。		
5	4	173	75				
6	5	178	83				
7	6	184	78				
8	7	189	95				
9	8	176	74				
10	9	183	85				
11	10	181	75				
12	11	184	96				
13	12	184	89				
14	13	174	77				
15	14	177	76				
16	15	173	70				
17	16	188	90				
18	17	182	79				
19	18	181	84				
20	19	180	86				
21	20	173	77				
22	21	178	80				
23	22	180	78				
24	23	180	87				
25	24	171	73				
26	25	178	81				
27	26	175	78				
28	27	170	66				
29	28	179	83				
30	29	174	73				
31	30	172	70				
32							

<発展編>

本書で紹介されていないテクニックを使うものもあります。適宜インターネットで検索しながら解いてみてください。

問題6 (Sheet 6)

あるクラスの 10 人の生徒の数学と英語の得点がまとめてあります。

右の図では、その合計点をもとにして、F 列に順位を表示させています。

しかし、7 位の生徒が 3 人いて、同率の順位が存在してしまっています。

そこで、以下の条件で「順位」が正しく表示できるようにこの表を改良してください。なお、表を改良するにあたって、必要であれば他の行や列を使ってもかまいません。

<条件>

- ・数学の得点が高い方を、より順位が高くなるようにする。
- ・数学の得点が同点の場合は、英語の得点が高い方を、より順位が高くなるようにする。
- ・数学も英語も同点の場合は、「番号」が若い方を、より順位が高くなるようにする。

	A	B	C	D	E	F
1	番号	氏名	数学	英語	合計	順位
2	1	ドラえもん	85	94	179	3
3	2	のび太	42	23	65	10
4	3	しずか	92	88	180	2
5	4	ジャイアン	42	55	97	7
6	5	スネ夫	68	29	97	7
7	6	出木杉	100	100	200	1
8	7	ドラミ	84	68	152	4
9	8	セワシ	42	55	97	7
10	9	ノビスケ	12	91	103	6
11	10	ジャイ子	56	82	138	5
12						

問題7 (Sheet 7)

ある果物農家が、商品の販売先と販売額をまとめた表をつくりました。

E2 セルに販売先、F2 セルに商品名を入力すると、G2 セルにその販売先に対する商品の販売額の合計が表示されるように、適切な関数を入力してください。

この例では、「上田青果」に「りんご」を販売した合計販売額 145000 円が表示されるようになります。

	A	B	C	D	E	F	G
1	販売先	商品名	販売額		販売先	商品名	販売額合計
2	井上商事	りんご	52,000		上田青果	りんご	
3	上田青果	りんご	65,000				
4	井上商事	みかん	78,000				
5	上田青果	みかん	43,000				
6	井上商事	りんご	45,000				
7	上田青果	バナナ	57,000				
8	井上商事	バナナ	49,000				
9	上田青果	りんご	80,000				
10							

問題8 (Sheet 8)

D 列に、数学の得点が 80 点以上ならば「優」、70 点以上ならば「良」、50 点以上ならば「可」、50 点未満ならば「不」が表示されるように適切な関数を入力してください。

	A	B	C	D
1	番号	氏名	数学	成績
2	1	ドラえもん	52	
3	2	のび太	74	
4	3	しずか	84	
5	4	ジャイアン	77	
6	5	スネ夫	68	
7	6	出木杉	59	
8	7	ドラミ	77	
9	8	セワシ	47	
10				