

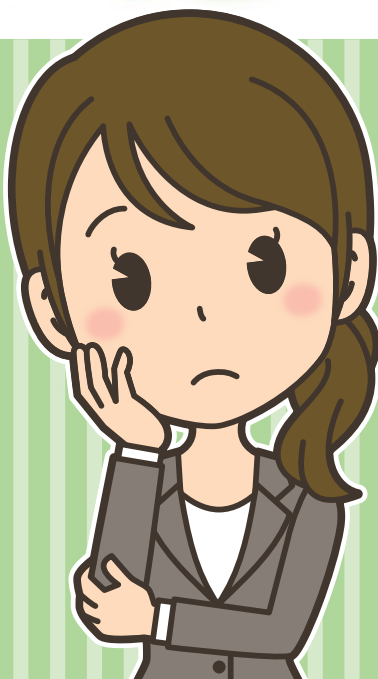
日文 教授用資料

中学校数学

データの活用

新教材の
指導の
手引き

累積度数



四分位範囲
箱ひげ図

本資料は、一般社団法人教科書協会「教科書発行者行動規範」に則り、配布を許可されているものです。

日文の実践事例、教科情報

詳しくはWebへ!

日文

検索

未来をになう子どもたちへ
日本文教出版

◆はじめに

平成 29 年 3 月に告示された新学習指導要領において、中学校数学科では、統計教育を重視するという方針が打ち出され、従来の「資料の活用」領域は「データの活用」領域に改称されました。

扱う内容・用語としては、「平均値、中央値、最頻値、階級」が中 1 から小 6 へ、「多数の観察や多数回の試行によって得られる確率」が中 2 から中 1 へそれぞれ移行するとともに、「累積度数」が中 1、「四分位範囲や箱ひげ図」が中 2 に加えられました。また、中学校ではすべての学年で「批判的に考察し」という文言が示されました。

さらに、『学習指導要領（平成 29 年告示）解説 数学編』（日本文教出版）には、「累積相対度数」や「相対度数を確率とみなすこと」についても述べられています。

本資料では、中学校数学科の「データの活用」領域における新教材を中心に、指導するにあたって知っておきたいことから、具体的な事例を挙げて解説しています。

既刊『小学校算数・中学校数学「データの活用」指導の初歩の初歩』（平成 30 年 5 月発行、日本文教出版ウェブページで公開中）と合わせて、先生方の教材研究の一助としていただければ幸いです。

◆もくじ

累積度数と累積相対度数【中 1】	2	2 年の練習問題(例)	18
統計的確率【中 1】	5	2 年の練習問題(例)の解答・解説	20
1 年の練習問題(例)	8	参考資料：累積相対度数と四分位数	21
1 年の練習問題(例)の解答・解説	9	批判的思考【中 1～中 3】	22
範囲と四分位範囲【中 1～中 2】	10	参考資料：移行措置での対応	24
箱ひげ図の基本【中 2】	12		
箱ひげ図の見方【中 2】	14		
ヒストグラムと箱ひげ図の対比【中 2】	16		

- ・本資料に示す指導学年は、新学習指導要領に基づくものです。
- ・本資料では、気象庁ウェブサイトより入手した気象データを取り上げています。そこに示されている「東京」「大阪」「福岡」「博多」などは、いずれも都市名ではなく気象データの観測地点名です。
- ・Microsoft® Excel® は、米国 Microsoft Corporation の米国およびその他の国における登録商標です。

累積度数と累積相対度数

新学習指導要領で、中1の〔用語〕に累積度数が追加されました。また、「累積相対度数」についても『学習指導要領（平成29年告示）解説 数学編』（日本文教出版）のp.90～91で言及されていて、その活用方法も示されています。

これらは移行措置の対象であり、平成31年度（令和元年度）以降は中1で指導します。

1. 累積度数

累積度数は、度数分布表で、最小の階級からある階級までの度数の合計のことです。

度数分布表では、右の表1（ある中学校の1年生40人の通学時間）のように表します。最大の階級までの累積度数は度数の合計（総度数）と

一致します。このような表を作成した際には必ず確認させましょう。

ポイント

上の表1において、10～15（10以上15未満）の階級までの累積度数に着目すると、通学時間が15分未満の生徒は11人いることがわかります。

また、表1からは、通学時間が短い方から10人目、20人目、30人目の生徒がそれぞれの階級に含まれるかを読み取ることができます。

2. 累積相対度数

累積相対度数は、度数分布表で、最小の階級からある階級までの相対度数の合計のことです。

表2は、表1に相対度数と累積相対度数の列を追加した表をExcelで作成したものです。また、表3は、表2で使用する関数や計算式を表示したものです。

表2

	A	B	C	D	E	F	G
1	時間(分)			度数 (人)	相対度数	累積度数 (人)	累積 相対度数
2	以上		未満				
3	0	～	5	1	0.025	1	0.025
4	5	～	10	3	0.075	4	0.1
5	10	～	15	7	0.175	11	0.275
6	15	～	20	9	0.225	20	0.5
7	20	～	25	12	0.3	32	0.8
8	25	～	30	8	0.2	40	1
9	合計			40	1		

表3

	A	B	C	D	E	F	G
1	時間(分)			度数 (人)	相対度数	累積度数 (人)	累積 相対度数
2	以上		未満				
3	0	～	5	1	=D3/\$D\$9	=D3	=E3
4	5	～	10	3	=D4/\$D\$9	=F3+D4	=G3+E4
5	10	～	15	7	=D5/\$D\$9	=F4+D5	=G4+E5
6	15	～	20	9	=D6/\$D\$9	=F5+D6	=G5+E6
7	20	～	25	12	=D7/\$D\$9	=F6+D7	=G6+E7
8	25	～	30	8	=D8/\$D\$9	=F7+D8	=G7+E8
9	合計			=SUM(D3:D8)	=SUM(E3:E8)		

ここではあえて、相対度数や累積相対度数を四捨五入していません。累積相対度数は、その定義の通り、最小の階級から当該の階級までの相対度数の合計として求めています。（次ページに示す㊦の求め方）

ポイント

累積相対度数から、ある階級までの全体に対する割合を知ることができます。

表2からは、40人のうち、5割の生徒の

通学時間は 20 分未満，8 割の生徒の通学時間は 25 分未満であることがわかります。

3. 累積相対度数の求め方

累積相対度数の求め方には，次の 2 通りがあります。

- ㊦ 各階級の相対度数を求めてから，当該の階級までの相対度数を合計する。
- ㊧ 当該の階級までの累積度数を求めてから総度数でわる。

総度数を N ，各階級の度数を最小の階級から順に n_1, n_2, n_3, \dots として，最小の階級から 3 つめの階級までの累積相対度数を式で表すと，それぞれ次のようになります。

$$\text{㊦} \quad \frac{n_1}{N} + \frac{n_2}{N} + \frac{n_3}{N} \qquad \text{㊧} \quad \frac{n_1 + n_2 + n_3}{N}$$

この 2 つの式は同じ数を表していることから，㊦ と ㊧ の求め方は，数学的には同義といえます。しかし，それぞれの方法で累積相対度数を求めた結果，双方の値にずれが生じる場合があります。

表 4 の累積相対度数は，㊦ の方法で求めています。各階級の相対度数は，小数第 3 位を四捨五入して，小数第 2 位までの近似値で表しています。

表 4 ㊦の方法で求めた場合

時間 (分)	度数 (人)	相対度数	累積相対度数
以上 未満			
0 ~ 5	1	0.03	0.03
5 ~ 10	3	0.08	0.11
10 ~ 15	7	0.18	0.29
15 ~ 20	9	0.23	0.52
20 ~ 25	12	0.30	0.82
25 ~ 30	8	0.20	1.02
合計	40	1.02	

最小の階級から各階級までの相対度数の合計

表 5 の累積相対度数は，㊧ の方法で求めています。累積相対度数は，表 4 の相対度数と同じようにして求めた近似値です。

表 5 ㊧の方法で求めた場合

時間 (分)	度数 (人)	累積度数 (人)	累積相対度数
以上 未満			
0 ~ 5	1	1	0.03
5 ~ 10	3	4	0.10
10 ~ 15	7	11	0.28
15 ~ 20	9	20	0.50
20 ~ 25	12	32	0.80
25 ~ 30	8	40	1.00
合計	40		

$(\text{累積度数}) \div (\text{総度数})$

表 4 と表 5 を比べると，累積相対度数の値に最大で 0.02 のずれが生じています。

㊦ は累積相対度数の定義通りであり，累積度数と同じ求め方であるため理解しやすいといえます。しかし，累積相対度数からデータの分布の傾向を読み取るという意味では，㊧ の方法で求めた値の方が，データの実態に合っているといえます。

具体例を挙げて考えてみましょう。

表 4 をもとに考えると，15 ~ 20 の階級までの累積相対度数が 0.52 であることから，「40 人のうち，過半数の生徒の通学時間は 20 分未満である」といえます。しかし，15 ~ 20 の階級までの累積相対度数が 20 人であることから，実際には「40 人の半分の 20 人の通学時間は 20 分未満である」が正しく，「過半数」は正しくないことがわかります。この点では，表 5 の累積相対度数 0.50 の方がデータの実態を表しているといえます。

ただし，㊦ の方法で求めた累積相対度数が間違いということではありません。「累積相対度数が 0.52 だから過半数といえるか」

といった問題は、誤差を含む可能性を読み手が考慮して判断するべきものだからです。

右の図1は、日本文教出版が発行している移行用補助教材です。ここで紹介している累積相対度数の求め方は、前ページに示した①です。新学習指導要領では、近似値や誤差、有効数字の内容が中3へ移行していることから、ここでは丸め誤差によるずれにつ

いては深入りを避け、①の方法のみを扱っています。

指導にあたっては、混乱を避ける意味で、⑦と①のどちらの方法で求めても結果が同じになるようなデータを教材とする方が無難といえます。

それでも、実在するデータや、自分たちで実験や調査をして集めたデータを分析する際には避けられない問題でもあるので、注意してください。

4. 相対度数の合計が1.00にならない場合の処理について

前ページの表4では、相対度数の合計が1.02となっています。このような場合、どのように処理するべきでしょうか。

小学校算数で、これと似た状況があります。小学5年では割合とともに円グラフや

(累積度数)
最小の階級から10分以上15分未満の階級までの度数の合計を求めます。
 $20+18+25=63$
よって、求める累積度数は63人です。

(累積相対度数)
累積相対度数は、累積度数を度数の合計でわると求められます。
 $\frac{63}{140}=0.45$
よって、求める累積相対度数は0.45です。

表2 通学時間(A中学校)

階級(分)	度数(人)	相対度数	累積度数(人)	累積相対度数
以上 未満 0 ~ 5	20	0.14	20	0.14
5 ~ 10	18	0.13	38	0.27
10 ~ 15	25	0.18	63	0.45
15 ~ 20	35	0.25	98	0.70
20 ~ 25	22	0.16	120	0.86
25 ~ 30	20	0.14	140	1.00
合計	140	1.00		

図1 中学数学 移行用補助教材 2019年度第1学年用/2020年度第2学年用 (日本文教出版)

備考 ・百分比は各問いの回答者数を100%として算出し、小数点第2位を四捨五入したため、百分比の合計が100%にならない場合がある。また、百分比の差を示す「ポイント」については、小数点第1位を四捨五入して示した。

図2 平成29年度「国語に関する世論調査」の結果の概要(文化庁ウェブサイト)

帯グラフを指導しますが、四捨五入で生じる丸め誤差の関係で割合の合計が100%にならないことがあります。しかし、円グラフや帯グラフをかく際、割合の合計が99%や101%になると、指導上の不都合が生じます。そこで、データの読み取りに影響が少なくなるよう、割合の一番大きい部分が「その他」で値を増減させて、全体が100%になるように指導するのが一般的です。

これにならって、中学校でも、相対度数の合計を1.00にするために、度数が最も大きい階級の相対度数を操作するという指導法があります。ただし、これらは、あくまでも小学生や中学生への指導上の配慮です。

実社会では、無理に1.00(100%)にするような調整をせずに、上の図2のように注意がきをするのが一般的です。

統計的確率

新学習指導要領で、「多数の観察や多数回の試行によって得られる確率」が中2から中1へ移行されました。いわゆる「統計的確率」の内容が中1へ移り、「数学的確率」の内容は中2に残ったといえます。

中1へ移った内容は移行措置の対象であり、令和2年度以降は中1で指導します。ただし、令和2年度までの中2には、従来通り指導します。

もう少し詳しく述べると、新学習指導要領では、中1で扱う確率の内容について、次のように示されています。

ア(ア) 多数の観察や多数回の試行によって得られる確率の必要性と意味を理解すること。
イ(ア) 多数の観察や多数回の試行の結果を基にして、不確定な事象の起こりやすさの傾向を読み取り表現すること。

ア(ア)は知識・技能に関する内容であり、イ(ア)は思考力・判断力・表現力等に関する内容です。このうち、イ(ア)で述べられていることは、後述する「相対度数を確率とみなすこと」を意味しています。

ここでは、中1で指導する確率について述べます。

1. 多数の観察や多数回の試行によって得られる確率

確率とは、あることがらの起こりやすさの程度を0以上1以下の数で表したもののことです。確率が p であるというのは、同じ観察や試行を多数回繰り返したとき、そ

のことがらが起こる相対度数が p に近づくという意味です。

相対度数は、全体に対する部分の割合を示す値です。多数の観察や多数回の試行においては、

$$\frac{\text{(ことがら A が起こった回数)}}{\text{(全体の回数)}}$$

が、ことがら A が起こった相対度数です。

例えば、ペットボトルのキャップを20回投げたときに表向きになった回数が4回なら、表向きの相対度数は0.2です。

図1は、1個のペットボトルのキャップを1000回投げるシミュレーションを3回行った結果です。図1で、20回投げた時点の表向きの相対度数は0.25、0.35、0.45とばらついていますが、投げる回数が多くなると、そのばらつきが小さくなっていくのがわかります。投げる回数をさらに増やしていくと、表向きの相対度数は、ある特定の値に近づいていきます。その値が、1個のペットボトルのキャップを1回投げたときに表が出る確率です。

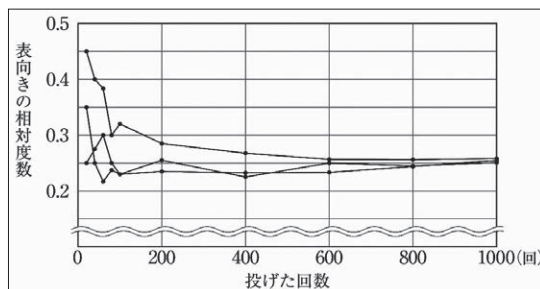


図1 ペットボトルキャップを1000回投げる実験をしたときの表向きの相対度数の変化の様子

15 次の(1)、(2)の各問いに答えなさい。

(1) 表と裏の出方が同様に確からしい硬貨があります。この硬貨を投げる実験を多数回くり返し、表の出る相対度数を調べます。このとき、相対度数の変化のようすについて、下のアからエまでのの中から正しいものを1つ選びなさい。

ア 硬貨を投げる回数が増えるにつれて、表の出る相対度数のばらつきは小さくなり、その値は1に近づく。

イ 硬貨を投げる回数が増えるにつれて、表の出る相対度数のばらつきは小さくなり、その値は0.5に近づく。

ウ 硬貨を投げる回数が増えても、表の出る相対度数のばらつきはなく、その値は0.5で一定である。

エ 硬貨を投げる回数が増えても、表の出る相対度数の値は大きくなったり小さくなったりして、一定の値には近づかない。

図2 平成30年度全国学力・学習状況調査A問題15(1)

2. 確率の意味の理解

上の図2は、平成30年度の全国学力・学習状況調査A問題です。正答はイ(相対度数は0.5に近づく)で、その正答率は40.2%です。他の反応率を見ると、ア(相対度数は1に近づく)が20.1%、ウ(相対度数は0.5で一定である)が15.8%、エ(相対度数は一定の値には近づかない)が22.2%、無解答が1.7%です。「確率の意味の理解」に課題があるといえるでしょう。

授業では、図1のような実験の結果と確率の意味を結びつけるような指導が重要です。理解をより確かなものにするためには、実際に実験をしてみることが有効と考えられます。

なお、「同様に確からしい」は中2で学習する内容なので、図2の問題を中1で扱う場合は、「投げたときに表が出る確率が0.5である硬貨があります。」とアレンジするとよいでしょう。

3. 相対度数を確率とみなすこと

相対度数は試行によって得られた結果、確定したことからについて表された数であ

るのに対して、確率は不確実なことから起こりやすさの程度を数で表したものです。つまり、相対度数と確率は同じものではありません。しかし、日常生活や社会において、相対度数を確率とみなして用いることがあります。

前述のイ(ア)について、『学習指導要領(平成29年告示)解説 数学編』(日本文教出版)では、図3を使った具体例を示して解説しています。

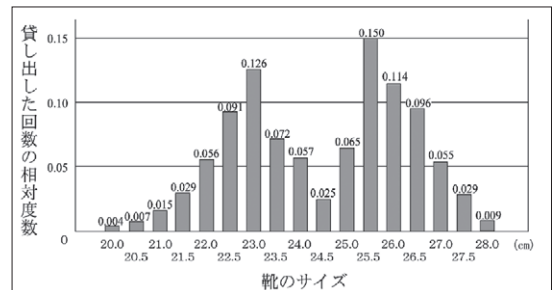


図3 『学習指導要領(平成29年告示)解説 数学編』p.93

上の図3は、あるボーリング場で過去1年間に貸し出された靴のサイズと、貸し出した回数の相対度数について表したものです。『解説』では、この相対度数を確率とみなして、それぞれのサイズの靴を何足買うか決めていくということを例として挙げています。

このように、新学習指導要領では、過去のデータから起こりやすさの傾向を予測するために、相対度数を確率とみなす活動を設定することを求めています。

相対度数を確率とみなして意思決定をする課題としては、平成24年度の全国学力・学習状況調査B問題3が参考になります。次ページの図4は、その問題の一部として提示されたヒストグラムです。

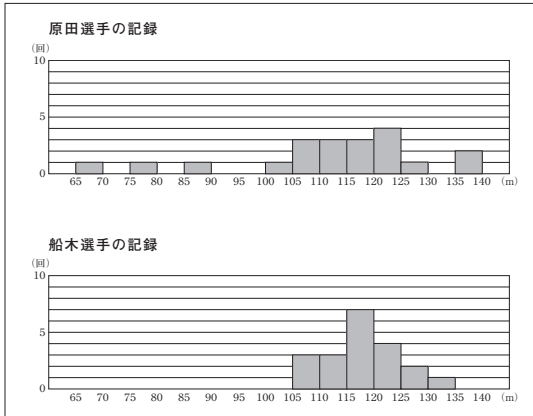


図4 平成24年度全国学力・学習状況調査B問題③の一部

原田選手が110m以上飛んだのは、20回のうち13回。



$13 \div 20 = 0.65$ だから、原田選手が110m以上飛んだ相対度数は0.65。



この相対度数を、「次の1回のジャンプで原田選手が110m以上飛ぶ確率」とみなすと、その確率は0.65。



同じように考えると、次の1回のジャンプで船木選手が110m以上飛ぶ確率は0.85。



次の1回でより遠くへ飛びそうな選手を1人選ぶとすると、110m以上飛ぶ確率が高い船木選手を選ぶ。

図4の2つのヒストグラムは、スキージャンプ競技の原田選手と船木選手の2人が長野オリンピック(1998年)までのいくつかの国際大会で飛んだ距離の記録をそれぞれまとめたものであり、総度数はどちらも20回です。小問(2)では、「2人のヒストグラムを比較して、次の1回でより遠くへ飛びそうな選手を1人選ぶとともに、選んだ理由を説明する」という趣旨の記述式の問題が出題されています。

この問題では、2つのヒストグラムの総度数が等しいことから、相対度数を使わず、度数で比較することもできますが、例えば「110m以上の階級の相対度数の合計で比べると、原田選手が($13 \div 20 =$)0.65、船木選手が($17 \div 20 =$)0.85で、船木選手の方が大きいから、船木選手を選ぶ。」と判断することが考えられます。これは、「110m以上の階級の相対度数の合計」を、「(それぞれの選手が次の1回のジャンプで)110m以上飛ぶ確率」とみなして意思決定の根拠としているということになります。

このとき、「次の1回のジャンプで原田選手が110m以上飛ぶ確率が0.65」の意味として適切なのは、次の㊶と㊷のどちらかを考えさせることで、確率は「必ず～になる」とは言い切れない不確実なことがらを考察の対象としていることを確認することも大切です。

- ㊶ 原田選手が100回ジャンプをすれば、そのうち65回は必ず110m以上飛ぶ。
- ㊷ 原田選手が100回ジャンプをすれば、そのうち65回くらいは110m以上飛ぶことが期待できる。

また、図4からは、船木選手が135m以上飛んだのは0回であることが読み取れます。このことをもとに、上のように相対度数を確率とみなして考えると、「次の1回のジャンプで船木選手が135m以上飛ぶ確率は0」となります。しかし、この場合、「船木選手が次のジャンプで135m以上飛ぶことはあり得ない」とは言い切れないという点に注意が必要です。

1 年の練習問題（例）

右の表1は、1951年から2018年までの68年間について、1年間の台風の発生数を表した度数分布表です。

表1から、例えば、1年間の台風の発生数が30個以上35個未満であった年が、この68年間に11回あったことがわかります。

表1をもとに、次の問いに答えましょう。
計算に電卓でんたくを使用しても構いません。

表1 台風の発生数（1951～2018年）

階級（個）	度数（回）	累積度数（回）	累積相対度数
以上 未満 10～15	1		
15～20	2		
20～25	25		
25～30	25		
30～35	11		
35～40	4		
合計	68		

（気象庁ウェブページのデータをもとに作成）

- (1) 各階級までの累積度数と累積相対度数を求めて、表1にかき入れましょう。
累積相対度数は、小数第3位を四捨五入して求めた概数がいすうで表しましょう。

- (2) この68年間で、台風の発生数が少ない方から10番目の年は、どの階級にふくまれていますか。

- (3) このデータの中央値は、どの階級にふくまれていますか。

- (4) 上の表をもとに考えると、来年の台風の発生数が25個未満である確率は約何%ということが出来ますか。一の位の数を四捨五入して、十の位までの概数で答えましょう。

1 年の練習問題（例）の解答・解説

(1) 表1 台風の発生数（1951～2018年）

階級（個）	度数（回）	累積度数（回）	累積相対度数
以上 未満 10～15	1	1	0.01
15～20	2	3	0.04
20～25	25	28	0.41
25～30	25	53	0.78
30～35	11	64	0.94
35～40	4	68	1.00
合計	68		

(2) 20個以上 25個未満（の階級）

(3) 25個以上 30個未満（の階級）

(4) 約40%

解説

(1) 累積度数と累積相対度数を求めることができるかを確認するための問題です。

累積相対度数の求め方は、p.3で述べたように、次の2通りの方法があります。

㊦ 各階級の相対度数を求めてから合計する。

㊧ 当該の階級までの累積度数を総度数でわる。

本問の場合、相対度数を求める必要がないことから、㊧の方法で求める方が合理的といえるでしょう。

また、本問では、㊦と㊧のどちらの方法で求めても累積相対度数の値がすべて一致しますが、データによっては、2通りの方法で求めた累積相対度数の値が一致しない場合があるので、問題を自作する際には注意してください。

(2) 累積度数に着目して、度数分布表から情報を読み取ることができるかを確認するための問題です。この場合、累積度数が初めて10回をこえるのがどの階級かを考えさせます。

つまりきが見られる生徒には、少ない方から1番目の年、2番目の年、3番目の年、…というように、順を追って考えさせるとよいでしょう。

(3) (2)と同様、累積度数に着目して、度数分布表から情報を読み取ることができるかを確認するための問題です。この場合、中央値が、台風の発生数が少ない方から数えて34番目と35番目の値の平均値であることを踏まえて、この両方の値が含まれる階級を答えとします。

なお、p.2の表1のように、度数分布表からでは中央値が含まれる階級を特定できない場合もあるので、問題を自作する際には注意してください。

(4) 累積相対度数に着目して、不確実な事象の起こりやすさの傾向について考察することができるかを確認するための問題です。この場合、累積相対度数は確率であるとはいえませんが、過去のデータから起こりやすさの傾向を予測するために、累積相対度数を確率とみなします。

このような内容を扱う際には、p.22で述べているように批判的思考を働かせて検討することも大切にしてほしいところです。

範囲と四分位範囲

四分位範囲と箱ひげ図が、高校数学Iから中2へ移行してきました。これらの内容は移行措置の対象であり、令和2年度以降は中2で指導します。ここでは、中1で指導する範囲（レンジ）と、中2で指導する四分位範囲について述べます。

1. 最小値、最大値と四分位数

はじめに必要な用語の定義をします。

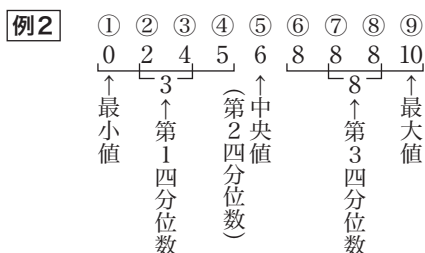
まず、データの個々の値を小さい順に並べ替えて、小さい方から大きい方に順番をつけます。最も小さい値が最小値、最も大きい値が最大値です。

次に、その順番について全体を4等分します。小さい方から4分の1のところの値が第1四分位数、4分の2（2分の1）のところの値が中央値（第2四分位数）、4分の3のところの値が第3四分位数です。第1四分位数、中央値、第3四分位数のいずれも、値がない場合は両隣の平均をとることで求めます。

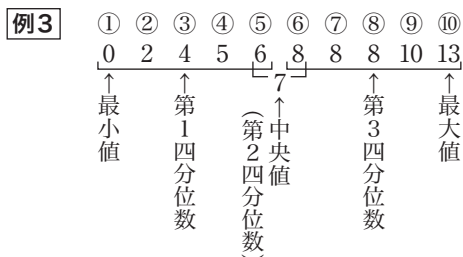
次の例1では、全部で11個の値があります。中央値は8です。第1四分位数は、中央値の1つ前までの5個の値の中央値4、第3四分位数は、中央値の1つ後からの5個のデータの中央値10となります。



次の例2では、全部で9個の値があります。中央値は6です。第1四分位数は、中央値の1つ前までの4個の値の中央値3（2と4の平均）、第3四分位数は、中央値の1つ後からの4個の値の中央値8（8と8の平均）となります。



次の例3では、全部で10個の値があります。中央値は7（6と8の平均）です。小さい方の半分（①～⑤）の値の中央値4が第1四分位数、大きい方の半分（⑥～⑩）の値の中央値8が第3四分位数となります。



データの最小値、第1四分位数、中央値、第3四分位数、最大値で分布の特徴を表すことを五数要約といいます。これを図に表したのが箱ひげ図です。

2. 範囲と四分位範囲

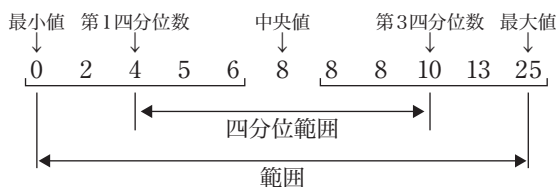
範囲（レンジ）とは最大値から最小値をひいた値、四分位範囲とは第3四分位数から第1四分位数をひいた値です。

$$(\text{範囲}) = (\text{最大値}) - (\text{最小値})$$

$$(\text{四分位範囲}) = (\text{第3四分位数}) - (\text{第1四分位数})$$

範囲は、データ全体の分布の広さを示します。それに対して四分位範囲は、中央値の前後の約25%ずつ、合わせて約50%のデータが入る区間の広さを示し、データが中央値付近にどのくらい散らばっているかを表します。四分位範囲が「小さい」とは約50%のデータが中央値付近に集まっていることを指し、四分位範囲が「大きい」とは約50%のデータが中央値付近に集まっていないことを指します。このように、四分位範囲は、データの散らばりの度合いを表す指標の1つとして用いられます。

前ページの例1のデータの範囲と四分位範囲を求めると、次のようになります。



$$(\text{範囲}) = (\text{最大値}) - (\text{最小値})$$

$$= 25 - 0$$

$$= 25$$

$$(\text{四分位範囲}) = (\text{第3四分位数}) - (\text{第1四分位数})$$

$$= 10 - 4$$

$$= 6$$

ここで、注意が必要なのは、四分位範囲が同じだからといって、中央値付近の分布が同じとは限らないということです。次の2つのデータを比べてみましょう。

A	0	1	2	3	3	11	12	13	14	15
B	0	1	2	6	7	7	8	13	14	15

どちらも範囲は15、四分位範囲は11、

中央値は7です。しかし、中央値付近の値の分布は全然違います。

それでも、範囲が同じで四分位範囲が違うデータ同士を比較する際には、四分位範囲が威力を発揮します。

3. 範囲と四分位範囲の特徴

範囲は、データの全ての値のうちで、最小値と最大値にあたる2つの値しか使いません。仮にデータの値が全部で1000個あっても、残りの998個の値が持つ情報は無視されます。また、範囲には、外れ値の影響を受けやすいことが欠点として挙げられます。

その点、四分位範囲はどうでしょうか。四分位範囲も、計算に使うのは第1四分位数、第3四分位数にあたる2つの値だけです。しかし、この2つの値を求める際には、全ての値を順に並べるという作業をします。このため、第1四分位数と第3四分位数には、中央値と同じように外れ値の影響を受けにくいという特徴があります。

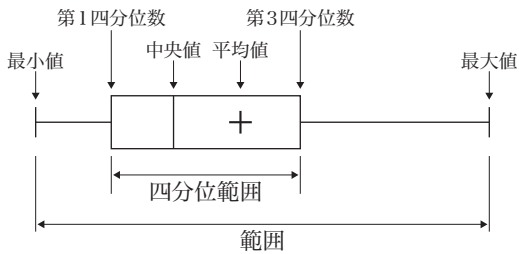
例えば、前ページの例3の最大値13を25に変えると、範囲は $(13 - 0 =) 13$ から $(25 - 0 =) 25$ に変わります。しかし、四分位数はいずれも変わらないので、四分位範囲も変わりません。

一般に、四分位範囲には、外れ値の影響を受けにくいという特徴があります。

箱ひげ図の基本

1. 箱ひげ図のかき方

箱ひげ図はデータの分布を視覚化するためのグラフであり、最小値、第1四分位数、中央値（第2四分位数）、第3四分位数、最大値を用いてかきます。

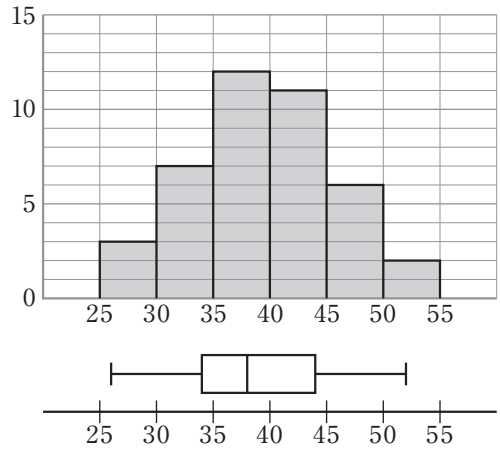


それぞれの値を示すところを縦線で表し、最小値と第1四分位数、第3四分位数と最大値は横線で結びます。第1四分位数と第3四分位数は箱形（長方形）に結びます。箱から出ている横線と最小値、最大値の縦線でできている図形が「ひげ」に見えるから「箱ひげ図」です。四分位範囲は、箱ひげ図の箱の長さに表示されます。箱ひげ図は横向きだけでなく、縦向きにかくこともあります。

必要があれば平均値を「+」でかき入れます。

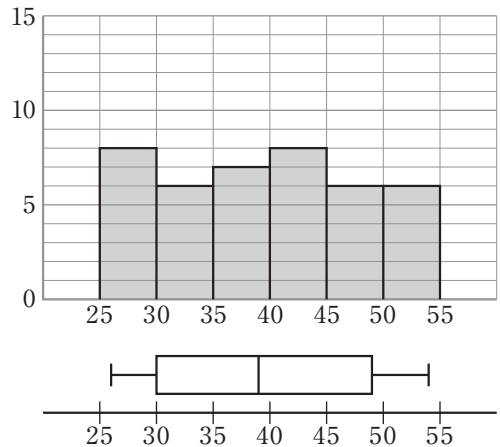
2. 箱ひげ図の基本的な見方

ヒストグラムが中央付近をピークとする急峻な山型になる分布のデータについて箱ひげ図をかくと、箱の長さが短くなります。箱の長さが短いということは四分位範囲が「小さい」ということであり、約50%のデータが中央値付近に集まっていることを指します。



中央にデータが集まっている例

ヒストグラムがなだらかな山型になる分布のデータについて箱ひげ図をかくと、箱の長さが長くなります。箱の長さが長いということは四分位範囲が「大きい」ということであり、約50%のデータが中央値付近に集まっていないことを指します。



全体にデータが散らばっている例

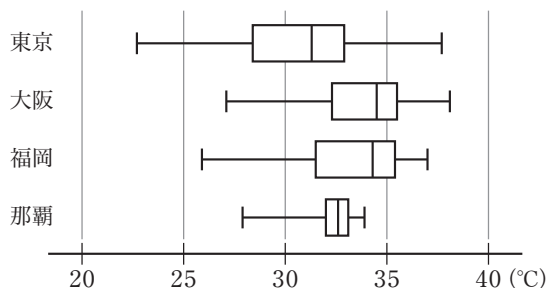
また、左右のひげが長いほど、その区間に含まれる値がばらけていて、ひげが短い

ほど、その区間に含まれる値が密集しているということがわかります。

3. 箱ひげ図の利点

箱ひげ図は、異なる複数のデータを比較するとき有効です。

次の図は4つの観測地点における2016年7月1日から8月31日までの62日間の最高気温のデータを箱ひげ図に表したものです。(データは気象庁ウェブサイトより)



この図を見ると、他の3観測地点に比べ、那覇のデータの範囲と四分位範囲が小さいことが一目でわかります。また、大阪と福岡を比べると、範囲と中央値、第3四分位数はほぼ同じですが、四分位範囲は福岡の方が大きいことなどがわかります。さらに、中央値の位置が箱の中で右に寄っていることや、左右のひげの長さからも、読み取ることができる情報があります。

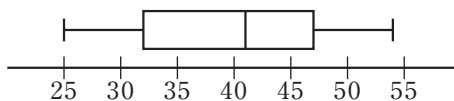
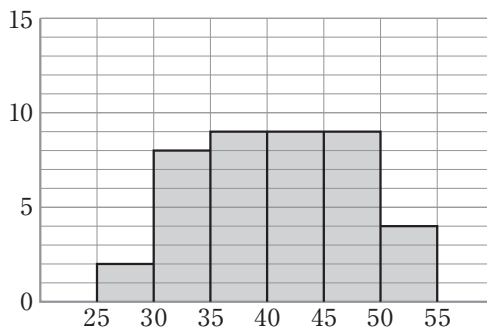
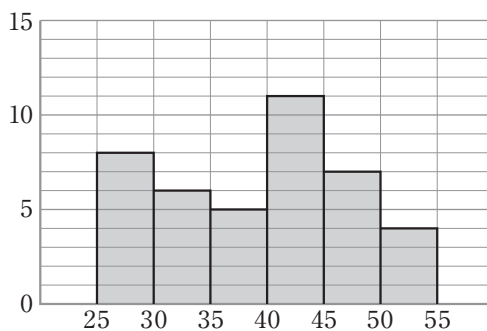
データの分布はヒストグラムでも見ることができますが、ヒストグラムでは情報が多い分、かえってデータの特徴をとらえにくい面があります。シンプルな箱ひげ図だからこそ見えてくることがあります。

4. 四分位範囲と箱ひげ図の注意点

簡潔さが箱ひげ図の長所といえますが、

反面、分布の形など、失われる情報もあります。

下に示す箱ひげ図は、最小値が25、第1四分位数が32、中央値が41、第3四分位数が47、最大値が54としてかかれています。次の2つのヒストグラムは形が異なりますが、いずれも下の箱ひげ図と対応しているものです。



この例からもわかるように、四分位範囲や箱ひげ図から一意に分布が決まるということはありません。そのため、四分位範囲や箱ひげ図を用いるときは、ヒストグラムと併用したり、平均値や最頻値と相補的に用いたりすることで、データの分布をよりの確にとらえ、表現できるようにする指導が求められます。

箱ひげ図の見方

箱ひげ図の特徴として、多数のデータの分布の傾向を比較しやすいという点が挙げられます。しかし、そのためには、まず、箱ひげ図のしくみを理解し、1つの箱ひげ図からデータの分布の情報を正しく読み取ることができることが前提となります。

ここでは、1つの箱ひげ図の見方の指導について述べます。

1. 1つの箱ひげ図の見方の指導

2018年の福岡の猛暑日（1日の最高気温が35℃以上の日）は16日でした。これは、例年に比べて「少ない」「ふつう」「多い」のどれといえるのでしょうか。

この疑問に答えるために表1を用意しました。これは、1998年から2017年までの20年間に、福岡で猛暑日は何日あったかを1年ごとに示したものです。

この20年間の福岡で、猛暑日が最も少なかったのは2009年で0日、最も多かったのは2013年で30日、このデータの範囲は30日です。

この場合、「16日は0～30日のおよそ真ん中だから『ふつう』といえる」という考えは妥当でしょうか。

「ふつう」という意味で使われる気象用語に「平年並み」があります。気象庁では、過去30年間（2019年現在では1981年～2010年を対象。2021年になると1991年～2020年というように10年ごとに更新される）の気象データの値30個を小さい順に並べ、小さい方から11番目から20番目までの値10個が分布する区間を「平年並み」と定義しています。

3等分か4等分かの違いはありますが、ここでは箱ひげ図の箱の部分を「ふつう」として考えてみることにします。つまり、箱ひげ図をかいて、その値が箱の部分にあてはまるのであれば「中央値付近の約50%にあたるから『ふつう』」、左のひげの部分にあてはまるのであれば「少ない方の約25%にあたるから『少ない』」、右のひげの部分にあてはまるのであれば「多い方の約25%にあたるから『多い』」と説明するのです。この「少ない」「ふつう」「多い」の基準は「平年並み」のように一般的に使われているものではありません。しかし、必要に応じて自分たちで基準を決めるということも大切です。

表1 年ごとの猛暑日の日数（1998～2017年、福岡）

年	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
日数(日)	5	3	5	8	2	2	6	3	6	8
年	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
日数(日)	10	0	24	8	10	30	2	4	22	13

（気象庁ウェブページのデータをもとに作成）

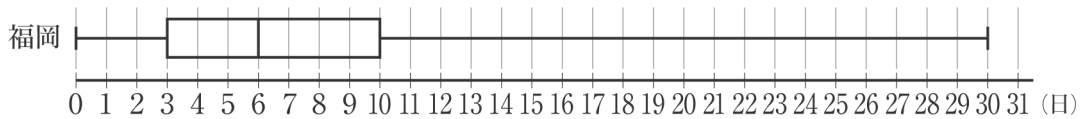


図1 年ごとの猛暑日の日数 (1998～2017年, 福岡)

図1は、表1のデータを箱ひげ図に表したものです。この図では、1年間の猛暑日が16日というのは右のひげの部分にあてはまるので、福岡の1年間の猛暑日としては「多い」といえます。

2. 箱ひげ図の見方の注意点

これまでに学んできた棒グラフやヒストグラムでは、棒(柱)の長さが対象となる数量に比例する関係にありました。このことから、図1のような箱ひげ図を見たとき、直観的に「左のひげは右のひげより短いから、左のひげの区間に分布する値の個数は右のひげの区間に分布する値の個数より少ない」とか「右のひげの長さは左のひげの長さの6倍以上あるから、右のひげの区間にある値の個数は左のひげの区間にある値の個数の6倍以上である」といった誤った解釈をしがちです。

そこで、箱ひげ図のしくみを指導する際には、図2のように、箱ひげ図とドットプロットを併記したのを見せ、「ひげや箱の長さが短いほど値の分布が密で、長いほどまばらである」ことや「ひげや箱が長くても短くても、4つに区切られた区間に分布する値の個数はおおよそ同じである」ことを印象づけるとよいでしょう。

また、図2を見ると、猛暑日が30日あった年が20年間で1回あったために右のひげが長くなっていることがはっきり認知でき、極端に離れた1つの値が箱ひげ図のひげの長さに影響を与えることも理解できます。その意味でも、学習の初期段階では、図2のように箱ひげ図とドットプロットを併記して見せることが、箱ひげ図の正しい見方の理解・定着に有効であると考えられます。

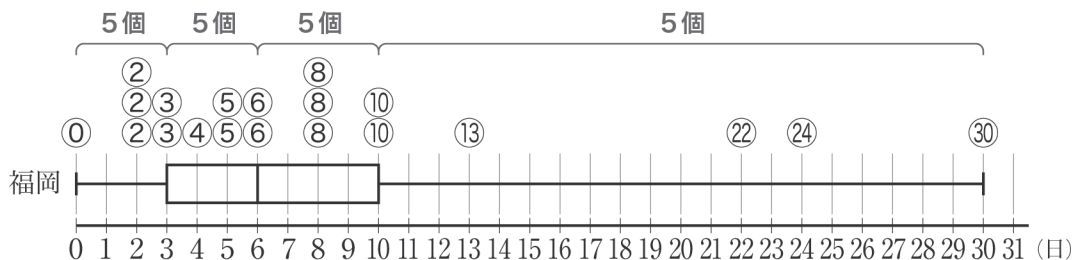


図2 図1に表1のデータのドットプロットを併記したもの

※図1、図2ともに『中学数学 移行用補助教材 2019年度第1学年用／2020年度第2学年用』（日本文教出版）より転載

ヒストグラムと箱ひげ図の対比

ヒストグラムと箱ひげ図から読み取ることができることについて、具体例を使って考えてみましょう。

次ページの図1の6つのヒストグラムは、2019年7月1日から8月31日までの、札幌、仙台、東京、大阪、博多、那覇における1日ごとの最高気温のデータをそれぞれまとめたものです。データは、気象庁ウェブページから入手したものです。

例えば、札幌では、この62日間で、最高気温が15°C以上20度未満だった日が2日あったことがわかります。

図2は、同じデータを箱ひげ図に表したものです。ただし、並べる順番を中央値の小さい順に並べかえ、博多と東京以外は地点名を伏せています。

図2の(A)～(D)がどの地点の図かを特定できるか考えてみましょう。

まず、(A)～(D)の箱ひげ図の左端に着目すると、(A)と(B)のデータの最小値は15°Cと20°Cの間にあり、(C)と(D)のデータの最小値は25°Cと30°Cの間にあることがわかります。したがって、(A)と(B)のデータには15°C以上20°C未満の階級に度数が1日以上あること、(C)と(D)のデータには25°C未満の2つの階級の度数が0日であることがわかります。

以上のことを踏まえて図1を見ると、(A)と(B)にあてはまるのは札幌と仙台、(C)と(D)にあてはまるのは大阪と那覇に特定されます。

ここで、(A)と(B)の箱ひげ図の右端に着目すると、(A)のデータの最大値は30°Cと35°Cの間にあり、(B)のデータの最大値は35°Cと40°Cの間にあることがわかります。したがって、(A)が札幌、(B)が仙台であることが確定します。

同じように、(C)と(D)の箱ひげ図の右端に着目すると、(C)のデータの最大値は30°Cと35°Cの間にあり、(D)のデータの最大値は35°Cと40°Cの間にあることから、(C)が那覇、(D)が大阪であることが確定します。

このような課題は、図から情報を読み取る力をのぼすための練習にもなります。

なお、図2の博多と東京の地点名を伏せた場合、図1と図2の対応関係を見抜くことはできません。なぜなら、図1のヒストグラムにおいて、それぞれのデータの最小値、四分位数、最大値を含む階級が、すべて一致しているからです。

p.18～19では、同じ図（ただし、図2は地点名を明かして順番も並べかえています）を使った練習問題（例）をご紹介します。複数の図から必要な情報を読み取る練習をするとともに、それぞれのグラフから読み取れることがらと読み取れないことがらについて考察させることもできます。

また、各地点のデータの分布の傾向を比較させて、どんなことがいえるかを話し合わせてもよいでしょう。

図1 1日ごとの最高気温 (2019年7月1日~8月31日)

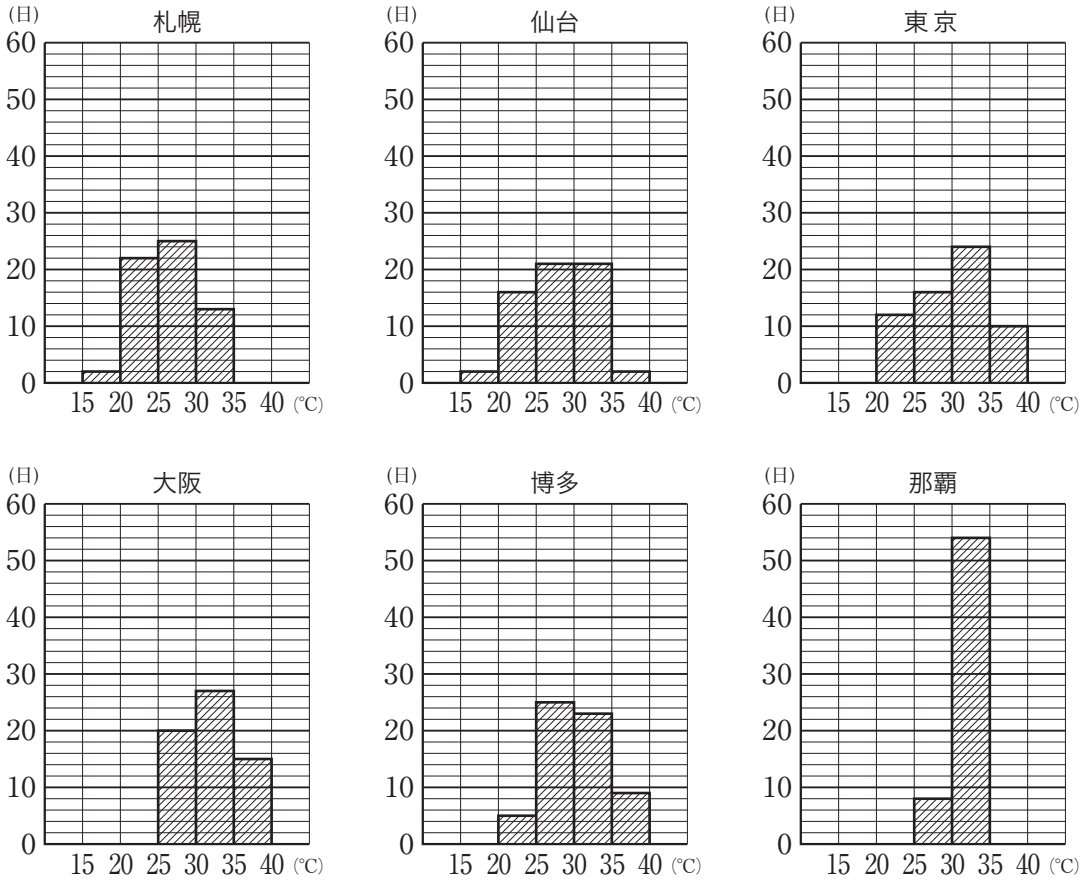
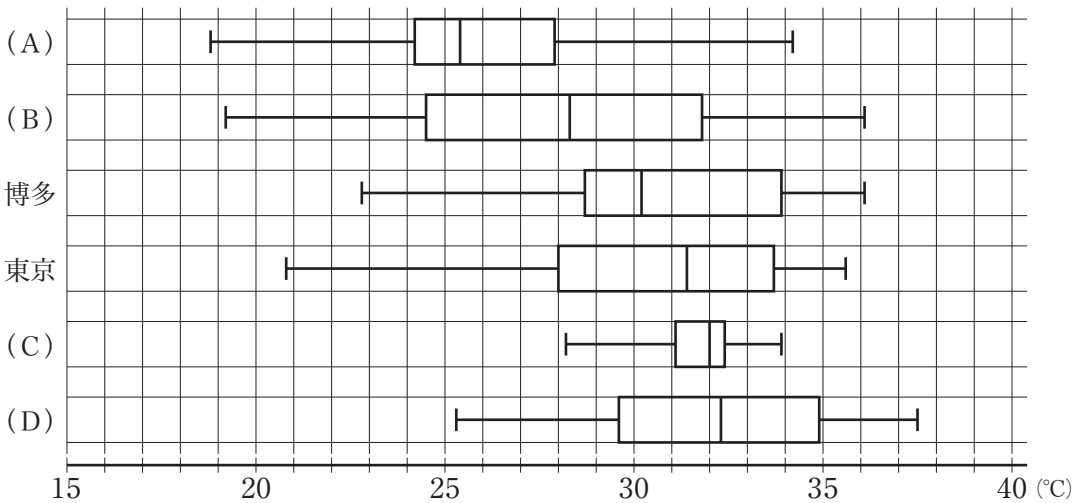


図2 1日ごとの最高気温 (2019年7月1日~8月31日)



2年の練習問題（例）

図1の6つのヒストグラムは、2019年7月1日から8月31日までの、6つの観測地点（札幌、仙台、東京、大阪、博多、那覇）における1日ごとの最高気温のデータをそれぞれまとめたものです。例えば、札幌では、この62日間で、最高気温が15°C以上20度未満だった日が2日あったことがわかります。

また、図2は、同じデータを箱ひげ図に表したものです。

図1のヒストグラムおよび図2の箱ひげ図について、次の問いに答えましょう。

- (1) 次の①～④の文章にあてはまる地点を、札幌、仙台、東京、大阪、博多、那覇の中から1つずつ選びましょう。
- ① 最小値が最も小さい地点 ② 中央値が最も大きい地点
③ 範囲が最も大きい地点 ④ 四分位範囲が最も小さい地点
- (2) 図1および図2から読み取ることができる内容として正しいものを、次の㉠～㉦の中からすべて選びましょう。
- ㉠ 札幌と那覇では、最高気温が35°C以上になった日がなかった。
㉡ 仙台では、最高気温が35°C以上になった日が1日だけあった。
㉢ 那覇の最小値より東京の第1四分位数の方が大きい。
㉣ 那覇の範囲より、仙台の四分位範囲の方が大きい。
㉤ 博多では、最高気温が28°C未満の日数が、35°C以上の日数の2倍以上あった。
㉥ 大阪は、四分位数の3つの値がすべて、6地点の中で最も大きい。
- (3) 6地点の中で、最高気温が35°C以上の日数が最も多かったのはどの地点ですか。また、このことを読み取ることができるのは、図1と図2のどちらですか。
- (4) 那覇では、この期間の半数以上の日の最高気温が31°C以上33度未満です。このことを読み取ることができるのは、図1と図2のどちらですか。

図1 1日ごとの最高気温 (2019年7月1日~8月31日)

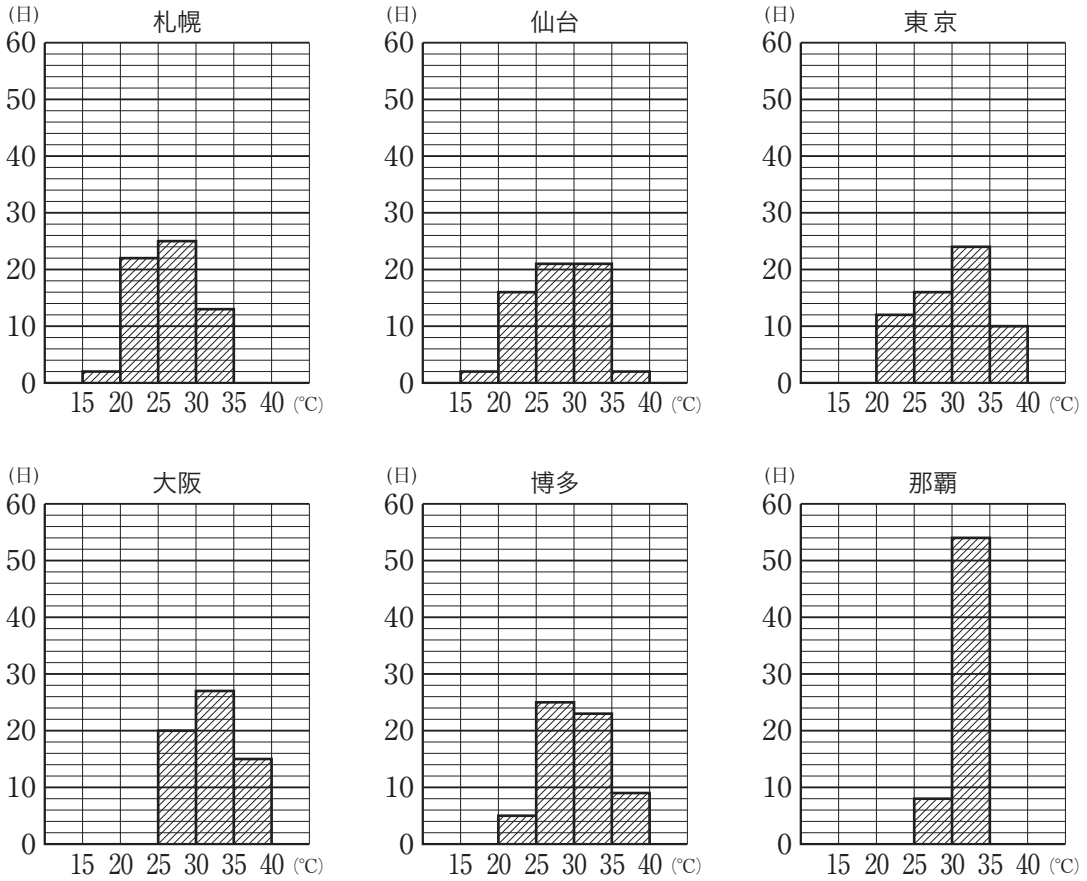
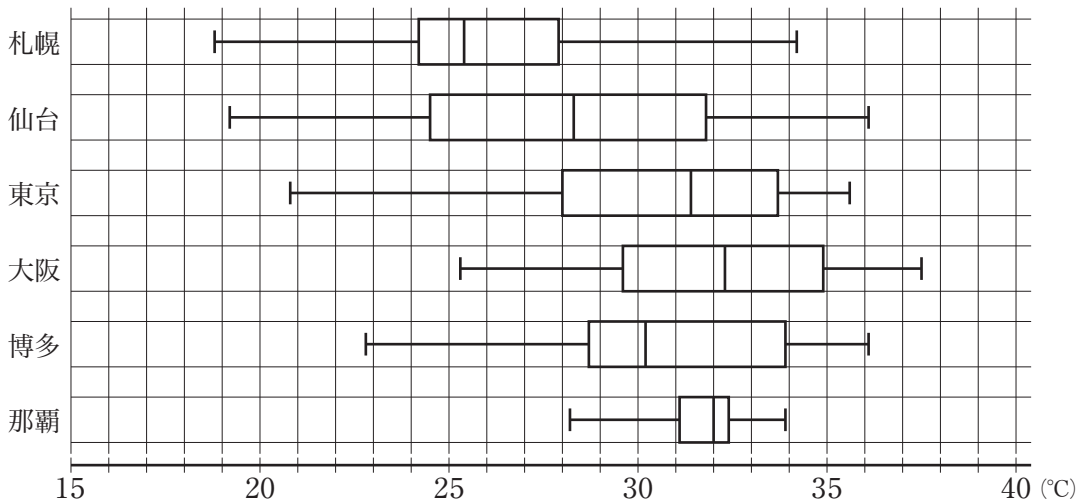


図2 1日ごとの最高気温 (2019年7月1日~8月31日)



2年の練習問題（例）の解答・解説

- (1) ① 札幌 ② 大阪
 ③ 仙台 ④ 那覇
- (2) ア, エ
- (3) 大阪, 図1
- (4) 図2

解説

(1) ①～④のいずれも、図2の箱ひげ図から基本的な情報を読み取ることができるかを確かめる問題です。

- ① 6つの箱ひげ図のうち、左端の位置が最も左側にあるものが答えです。
- ② 6つの箱ひげ図のうち、中央値を表す縦線が最も右側にあるものが答えです。
- ③ 6つの箱ひげ図のうち、全体の長さが最も長いものが答えです。
- ④ 6つの箱ひげ図のうち、箱の部分の長さが最も短いものが答えです。

(2) 図1のヒストグラムと図2の箱ひげ図から、必要な情報を選択して読み取ることができるかを確かめる問題です。

- ア 図1, 図2のどちらからでも、正しいことを読み取ることができます。
- イ 図1から、正しくないことを読み取ることができます。図1の縦軸の1めもりが度数の2日を表していることに注意が必要です。
- ウ 図2から、正しくないことを読み取ることができます。
- エ 図2から、正しいことを読み取るこ

とができます。那覇の範囲は 6°C より小さく、仙台の四分位範囲は 6°C より大きいことが(7°C より大きいことも)読み取ることができます。

- オ 図1から、博多の 35°C 以上の日数は9日とわかります。その2倍以上ということは、18日以上ということです。一方、図2から、博多の 28°C 未満の日数は15日以下であることがわかります(値の総数が62個だから、第1四分位数は小さい方から16番目の値と一致する)。

以上より、正しくないことを読み取ることができます。

- カ 図2から、正しくないことを読み取ることができます。

中央値と第3四分位数は大阪が最も大きいのですが、第1四分位数是那覇の方が大きいことが読み取れます。

- (3) ヒストグラムからは読み取ることができて、箱ひげ図からは読み取ることができないことがらです。

- (4) 箱ひげ図からは読み取ることができて、ヒストグラムからは読み取ることができないことがらです。那覇のヒストグラムからは、 30°C の日が54日あった可能性も否定できません。一方、箱ひげ図からは、 31°C から 33°C の間の区間に箱が収まっていることから正しいことがわかります。

参考資料：累積相対度数と四分位数

下の表1は、ある地区の中学1年生男子のハンドボール投げの記録を整理した度数分布表です。そのとなりの図は、表1の階級を横軸にとり、累積相対度数を縦軸にとってかいた折れ線グラフです。このようなグラフを、累積相対度数折れ線、累積相対度数グラフ、累積分布図などによびます。

このグラフで、例えば、縦軸の0.25のところから横に見て、折れ線にぶつかったところから下へ見ると、横軸の15 mのところからぶつかります。このことから、「この地区の中学1年生男子の約25%は、ハンドボール投げの記録が15 m未満である」と読み取ることができます。これは、データの第1四分位数が約15 mであることを意味します。

同じように、縦軸の0.50や0.75のところから見ていくと、データの中央値（第2四分位数）は約19 m、第3四分位数は約23 mであることを読み取ることができます。

この場合、各階級に含まれている測定値の値が不明なので、正確な四分位数は求められません。しかし、累積相対度数折れ線を使うと、四分位数がどのあたりにあるかをとらえることができます。

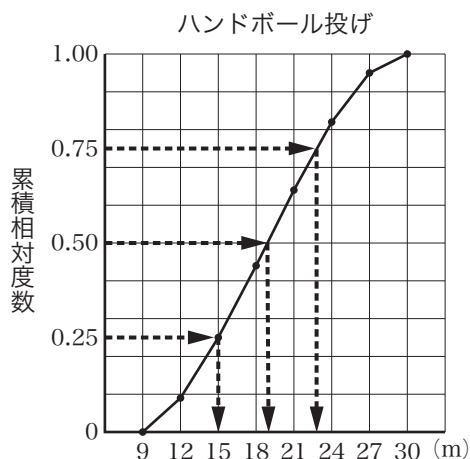
また、累積相対度数が0.90のところに着目すれば、「この地区の中学1年生男子の約90%は、ハンドボール投げの記録が26 m未満である」と読み取ることができます。このとき、「このデータの90%点（パーセント点）は約26 mである」といいます。

一般に、データの値を小さい順に並べ替えたとき、小さい方から $x\%$ の位置にある値を、 x パーセント点、または x パーセントイルといいます。第1四分位数、中央値、第3四分位数のことを、それぞれ25%点、50%点、75%点ともいいます。

※このページの内容は、学習指導要領上、中学校数学科で扱うと決められたものではありません。

表1 ハンドボール投げ

階級 (個)	度数 (回)	累積度数 (回)	累積相対度数
以上 未満 9 ~ 12	18	18	0.09
12 ~ 15	33	51	0.25
15 ~ 18	40	91	0.44
18 ~ 21	42	133	0.64
21 ~ 24	38	171	0.82
24 ~ 27	27	198	0.95
27 ~ 30	10	208	1.00
合計	208		



新学習指導要領では、中学校数学科の全学年で「批判的に考察し」という文言が入りました。小学校算数科でも6年で「批判的に考察すること」という文言が入っており、義務教育において重要な内容と位置づけられています。

ここでは、「批判的思考」について、具体例を挙げて述べます。

1. 根拠となるデータの妥当性について批判的に考察すること

p.8の「1年の練習問題(例)」の(4)では、過去(1951年から2018年までの68年間)のデータをもとに、「来年の台風の発生数が25個未満である確率は約40%」と結論づけています。

しかし、近年は異常気象ともよべる現象が多発しています。記憶に新しいところでも台風が日本列島に大きな被害をもたらしていることから、台風の発生数は増加する傾向にあるのではないかと推測することができます。そうすると、「過去のデータの傾向をもとに将来を予想する」という考え方が使えるのか疑わしくなり、上記の結論を見直す必要性も出てきます。

そこで、p.8の表1のもとになったデータにさかのぼり、1951年から1984年までの34年間と1985年から2018年までの34年間に2分割して、その2つのデータの分布を比較し、分布の傾向に変化があるかどうかを調べることで、上記の結論の妥当性を検討してみたいと思います。

台風の発生数(1951～2018年)

階級(個)	度数(回)	
	1951～1984年	1985～2018年
以上 未満 10～15	0	1
15～20	1	1
20～25	13	12
25～30	10	15
30～35	7	4
35～40	3	1
合計	34	34

(気象庁ウェブページのデータをもとに作成)

上の度数分布表で、20個以上25個未満の階級までの累積度数は、どちらのデータでも14回で等しく、その累積相対度数を求めると0.41となります。よって、前述の結論を否定するような変化は見られないことから、この結論は妥当といえることができます。

2. データの分析の仕方や着目する点について批判的に考察すること

上の表で30回以上の2つの階級の度数の合計に着目すると、1951年から1984年までの34年間では10回、1985年から2018年までの34年間では5回と、半減しています。そこで、今度は「台風の発生数は減少する傾向にあるか」という新たな問題を設定し、考察してみたいと思います。

次の図1は、1951年から2018年までの68年間における台風の発生数のデータを折れ線グラフに表したものです。この図

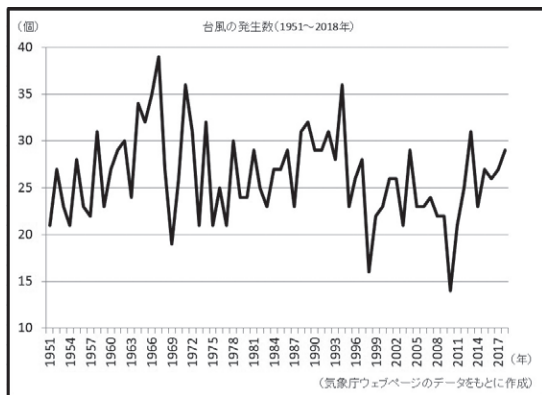


図1

を見ると、台風の発生数について、少なくとも増加傾向にあるという印象は受けません。むしろ、減少しているように見えませんか。

そこで、今度は前ページの度数分布表をもとに、度数分布多角形をかいてみることにします (図2)。

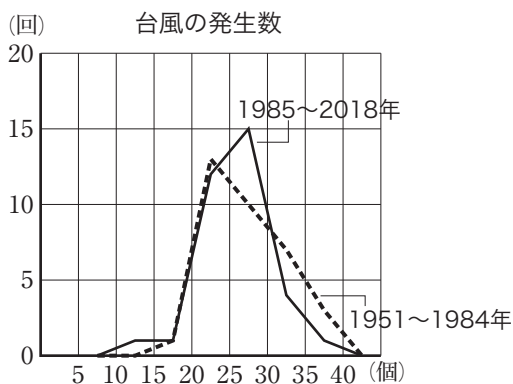


図2

山のピーク、すなわち最頻値に着目すると、1951～1984年より1985～2018年の34年間の方が大きいことから、台風の発生数は増加傾向にあるといえるかもしれません。しかし、発生数が15個未満の階級や30個以上の階級に着目すると、逆に減少傾向にあるということもできます。

さらに、図2と同じデータの分布を箱ひげ図に表してみます (図3)。

図3では、最小値と最大値がそれぞれ減少傾向にあることから、「台風の発生数が減少傾向にある」という印象を受けるかもしれません。しかし、箱の位置と長さ、中央値に着

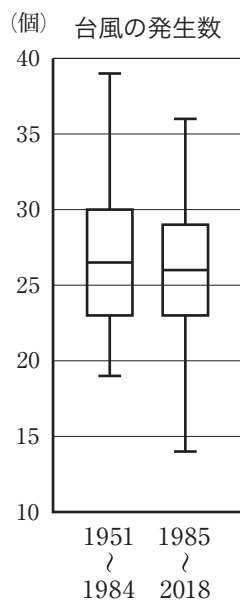


図3

目すると、「あまり変わっていない」ということもできそうです。

以上のことを総合的に判断すれば、これらのデータからは、「特に増加や減少の傾向は読み取ることができない」という結論が妥当ではないかと思われます。

3. 批判的に考察しようとする態度の育成

直観を頼りに仮説を立てること自体は悪いことではありませんが、その正しさの検証をすることなく、仮説が正しいと思込むことは危険です。

「データの活用」領域における学習では、データの収集の仕方は適切か、どの代表値が根拠としてふさわしいか、分布の形に着目しているか、分析した結果から得られる結論が妥当かなど、解決の過程を振り返って検討し、よりよい解決や結論を見いだそうとする態度を養うことを重視した指導が望まれます。

参考資料：移行措置での対応

■中学校数学科 移行措置の内容一覧（全領域）

平成 31 年度（令和元年度）	令和 2 年度
<p>(中 1 で追加する内容)</p> <ul style="list-style-type: none">・自然数を素数の積として表すこと・累積度数	<p>(中 1 で追加する内容)</p> <ul style="list-style-type: none">・自然数を素数の積として表すこと・累積度数・多数の観察や多数回の試行によって得られる確率
<p>(中 1 で省略する内容)</p> <ul style="list-style-type: none">・誤差や近似値，$a \times 10^n$ の形の表現 →中 3 で指導する。	<p>(中 1 で省略する内容)</p> <ul style="list-style-type: none">・誤差や近似値，$a \times 10^n$ の形の表現 →中 3 で指導する。
	<p>(中 2 で追加する内容)</p> <ul style="list-style-type: none">・四分位範囲や箱ひげ図



中学校数学

「データの活用」新教材の指導の手引き

日文 教授用資料

令和元年（2019年）11月29日発行

編集・発行人 佐々木秀樹

発行所 日本文教出版株式会社

〒558-0041 大阪市住吉区南住吉4-7-5

TEL: 06-6692-1261

本書の無断転載・複製を禁じます。

CD 33488

日本文教出版 株式会社
<https://www.nichibun-g.co.jp/>

大阪本社 〒558-0041 大阪市住吉区南住吉4-7-5
TEL:06-6692-1261 FAX:06-6606-5171

東京本社 〒165-0026 東京都中野区新井1-2-16
TEL:03-3389-4611 FAX:03-3389-4618

九州支社 〒810-0022 福岡市中央区薬院3-11-14
TEL:092-531-7696 FAX:092-521-3938

東海支社 〒461-0004 名古屋市東区葵1-13-18-7F・B
TEL:052-979-7260 FAX:052-979-7261

北海道出張所 〒001-0909 札幌市北区新琴似9-12-1-1
TEL:011-764-1201 FAX:011-764-0690