

データの分布の傾向をとらえよう

—景気の変化を家計調査のデータからとらえるには—

【分布の形・代表値（平均値・中央値・最頻値）・ヒストグラム・度数分布多角形】

1. 問題について

現在、「景気が悪い」という言葉をよく耳にします。そして、「収入減」という言葉もよく聞かれます。これは日本全国のすべての収入層、すなわち、収入の多い億万長者から、生活に困窮している人のすべての収入が減っているということなのでしょうか。総務省の家計調査の年間平均の所得金額のデータをもとに考えてみましょう。

総務省統計局では、年間平均の所得金額のデータを処理するときに、年間収入の大きさによって区分して表している。すべての世帯を収入の低い方から高い方へと順番に並べ、それを世帯数の上で10等分して10のグループを作り、収入の低いグループから順に、第1, 第2, 第3, …第10分位階級と呼んでいる。それぞれの階級について収入と支出をまとめたものが年間収入十分位階級別の結果である。所得水準は消費支出や消費支出の構成に影響する最も基本的な要因であるため、家計分析を行う上で欠くことのできない結果表である。なお、この年間収入十分位階級別結果のほかに、5等分して5つのグループを作る年間収入五分位階級別の結果も作成している。」総務省統計局「家計調査のしくみ」より (<http://www.stat.go.jp/data/kakei/10.htm>)

具体的に、1000件の世帯数で区切っていくと、収入は以下のように区分される。

区分	1	2	3	4	5	6	7	8	9	10
下限	0~	190~	269~	332~	393~	461~	546~	641~	774~	991~
上限	190	269	332	393	461	546	641	774	991	
世帯数	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000

この収入で区切って、各区分につき、収入1万円あたりの世帯数（密度）を算出し、その割合を表すと以下の右図のようなヒストグラムがかかる。（収入は単位：万円）

区分	1	2	3	4	5	6	7	8	9	10
下限	0~	190~	269~	332~	393~	461~	546~	641~	774~	991~
上限	190	269	332	393	461	546	641	774	991	*1705
世帯数	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
階級幅	190	79	63	61	68	85	95	113	217	714
密度 (世帯数割合)	5.263	12.658	15.873	16.393	14.706	11.765	10.526	7.519	4.608	1.401

※上限値は不明なので下限値+（平均-下限値）×2 なお、平均値は、547（万円）

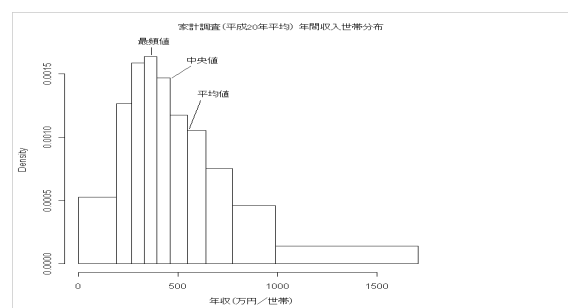
これからわかるように、収入の平均値、中央値、最頻値は一致しない。

他の年度を求めてもこの傾向は変わらず、

「最頻値<中央値<平均値」の関係は変わらない。

また、最頻値ごとの推移、中央値ごとの推移、平均値の推移をそれぞれ年度ごとに比較しても年々減少していることがわかる。これから、収入の傾向は毎年、やや下の方に集中して分布していることがわかる。

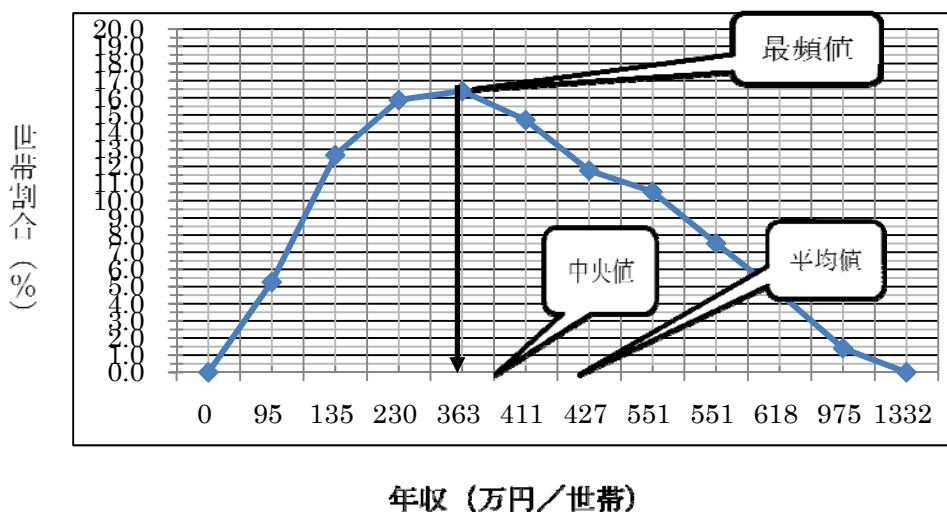
平成20年度世帯分布のヒストグラム



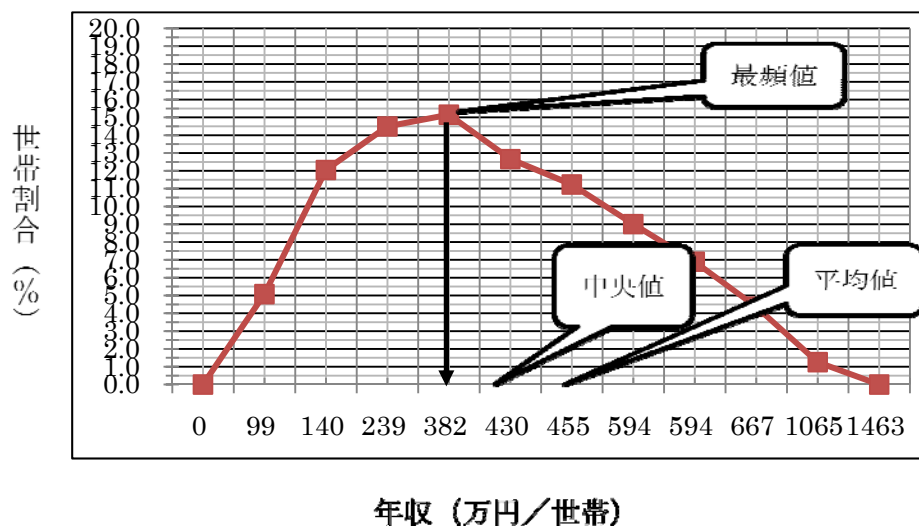
これからも景気後退の減少は、どの収入層の人にも関係しており、平均収入が中間層よりもやや高めと感じる人が多いと思われる。

年度ごとの比較をするのに、度数分布多角形を用いると推移の様子がよくみえる。よって、ヒストグラムを度数分布多角形で表してみる。

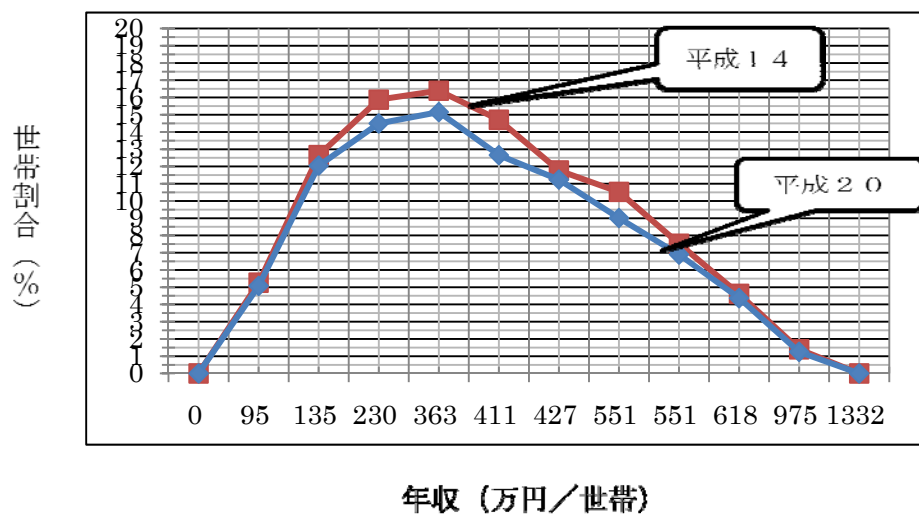
平成20年度の度数分布多角形



平成14年度の度数分布多角形

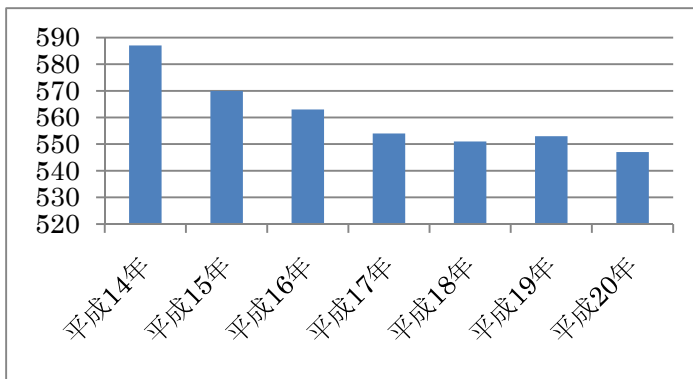


平成20年と14年の度数分布多角形



データの傾向を比較する観点として、平均値のみで捉える傾向にある。年ごとの傾向をより詳細に比較するためには、平均値の比較でよいのかという観点で考えさせる。

総世帯（二人以上の世帯+単身世帯）の年間収入の平均値の推移



左図は平成14年から平成20年の平均値の推移を棒グラフで表したものである。これから、平均値で比較した場合、年々、収入は減少傾向にあることがわかる。

一般に、平均でデータの傾向を捉えている傾向にあるが、家計調査の場合、家計収支を所得階層別に分類しているということは、平均値による比較は可能であるが、データの傾向を中央値、最頻値で捉える必要もあるということである。年次推移、および収入の傾向を考えていくにあたって、分布の特徴を考えていく必要があり、

総務省統計局では、代表値として平均値を代表値として用いている。ホームページにはその理由が以下のように記されている。

「家計調査の結果は、調査世帯数などを除けば平均値の結果である。どうして平均値を使うのかということを考えてみる。例えば、ここに平均値ではなく個々の世帯の実収入をそのまま一つ残らず網羅した表があるとする。しかし、このままの表から実収入について何かの判断を得るということは、非常に困難なことである。そこで、この表の内容を代表する値（代表値）を工夫することが必要となる。この代表値としては、算術平均値のほか最頻値、中位数、幾何平均値などがあり、通常、算術平均値を用いる。これは、算術平均値が分析する場合に計算が容易で理解し易く、また数学的に便利な点があるからである。しかし、算術平均値にも問題がないわけではない。そこで、この平均値の性質について家計調査の結果を例に考えてみる。

平均値はあくまでも集団全体を一つの数値で代表させたものであるが、例えば、全国・勤労者世帯の平成16年4月の実収入は1世帯当たり平均で487,026円で、この平均値以上の世帯数と平均値以下の世帯数は半々でなく、全体の約55.8%の世帯が平均値よりも低くなっている。世帯数を二等分する中位数は455,560円となっています。このように中位数が平均値を下回る場合平均値は実感よりも高めに感じられる傾向がある。前述とは逆に、平均値が個々の実感より低めに計算される例を述べてみる。例えば、全国・全世帯の平成16年4月の家賃地代は1世帯当たり平均で11,991円となっている。しかし、民間の借家に住んでいる世帯にとっては大変低く感じられる。これは、持家や家賃の低い給与住宅に住んでいる世帯も含めた全部の世帯の平均だからです。民間の借家の家賃をみるには、住居の所有関係別の結果表をみる必要がある。このような関係は家賃に限らず、教育費、電気冷蔵庫の購入金額など数多くある。学校へ行っている世帯員がいる世帯、電気冷蔵庫を購入した世帯にとっては、それらの数字はかなり低いと感じるのは以上のような理由から生じている。平均値は、以上のような留意点がありますが、家計調査の本来の目的であるグループ相互の比較や月々、年々の比較には、欠くことのできないものである。」総務省統計局「家計調査のしくみ」より

<http://www.stat.go.jp/data/kakei/10.htm>

このように、家計調査における世帯収入のデータの特徴を捉える代表値として、左に偏った分布では一般に中央値を代表値として用いるが、総務省統計局では平均値のメリットを用いて比較している。

このデータを用いて、平均値、中央値、最頻値の3つの関係を考えていくことも分布の傾向を捉える上で重要なことであることをヒストグラム、度数分布多角形をもとに捉えさせていきたい。

## 2. 授業について

### (1) 授業計画

中学校 1 年生「資料の活用」の代表値を学習した後に扱う。

第 1 時 家計調査の意味と年度毎の比較について考える。

第 2 時 家計調査の傾向についてどのようなことがわかるか、

### (2) 授業目標

- ・ 3 つの代表値が異なる場合があることを知り、そのデータの傾向をつかんでいく。
- ・ 代表値をヒストグラム・度数分布多角形と関連させてとらえていく。

### (3) 授業展開

	主な発問	予想される反応	留意点
導入	<p>「現在、“景気が悪い”という言葉をよく耳にします。また、“収入減”という言葉もよく聞かれます。これは日本全国、収入の多い億万長者から、生活に困窮している人すべての階層の人の収入が減っているということなののでしょうか。」</p> <p>「では、このことについて、総務省の家計調査の年間平均の所得金額のデータをもとに考えてみましょう。」</p> <p>「総務省の家計調査の結果によると、例えば、全国・勤労者世帯の平成 16 年 4 月の実収入は 1 世帯当たり平均で 487,026 円で、この平均値以上の世帯数と平均値以下の世帯数は半々でなく、全体の約 55.8% の世帯が平均値よりも低くなっており、世帯数を二等分する中央値は 455,560 円となっていると記されていますが、これから平成 16 年の収入についてどのようなことがわかりますか。」</p> <p>「このように中央値が平均値を下回る場合、平均値は実感よりも高めに感じられる傾向があります。」</p>	<p>「収入の高い人は収入が安定しているように思うので、あまり関係がないのではないか。」</p> <p>「すべての国民の平均的な収入が減少していると思う。」</p> <p>「”派遣切り”とか、”ネット難民”というように、収入が平均以下の人の収入が減っていると思う。」</p> <p>「収入が低い人の方が高い人よりも多いから、中央値&lt;平均値となるのではないか。」</p> <p>「平成 16 年の収入の分布は、収入が低い人が多い分布なのではないか。」</p> <p>「半数以上の人の方が平均値よりも低いのであるから、平均収入以下の人が半数以上いる。やはり、低所得者が多い。」</p>	<p>必要に応じて新聞記事などを提示。</p> <p>代表値（平均値、中央値、最頻値）が一致しない場合、どのような分布の傾向が考えられるか、意見を出させる。</p>
展開	<p>「では実際にデータをみてみよう。」</p> <p>総務省統計局では、年間平均の所得金額のデータを処理するときに、年間収入の大きさによって区分して表しています。すべての世帯を収入の低い方から高い方へと順番に並べ、それを世帯数の上で 10 等分して 10 のグループを作り、収入の</p>		<p>まず、総務省のデータの表示方法について説明する。このように表現することにより、収入の階層別の</p>

低いグループから順に、第1、第2、第3、…第10分位階級と呼んでいます。それぞれの階級について収入と支出をまとめたものが年間収入十分位階級別の結果です。所得水準は消費支出や消費支出の構成に影響する最も基本的な要因であるため、家計分析を行う上で欠くことのできない結果表です。なお、この年間収入十分位階級別結果のほかに、年間収入五分位階級別の結果も作成しています。」

「以下の表は、収入により世帯数を1000件で区切ったときの上限と下限を示したものです。」

データが得られることを理解させる。

データとヒストグラムを提示。

具体的に、1000件の世帯数で区切っていくと、収入は以下のように区分される。

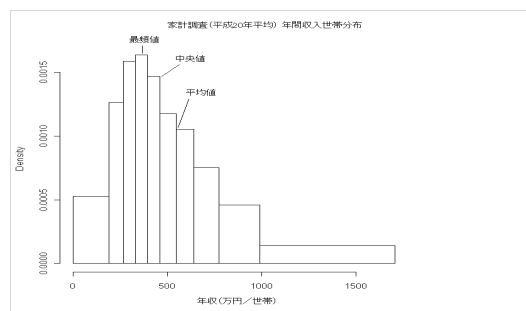
区分	1	2	3	4	5	6	7	8	9	10
下限	0~	190~	269~	332~	393~	461~	546~	641~	774~	991~
上限	190	269	332	393	461	546	641	774	991	
世帯数	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000

これから、階級幅と世帯数割合を算出すると以下ようになる。

区分	1	2	3	4	5	6	7	8	9	10
下限	0~	190~	269~	332~	393~	461~	546~	641~	774~	991~
上限	190	269	332	393	461	546	641	774	991	
世帯数	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
階級幅	190	79	63	61	68	85	95	113	217	714
密度 (世帯数割合)	5.263	12.658	15.873	16.393	14.706	11.765	10.526	7.519	4.608	1.401

「この分布から、どのようなデータの特徴がわかりますか。」

平成20年度世帯分布のヒストグラム



「最頻値<中央値<平均値」の関係がある。」

「収入が低い層の人が多く分布している左に偏った分布。」

「総務省の家計調査の結果によると、例えば、全国・勤労者世帯の平成16年4月の実収入は1世

<p>帯当たり平均で487,026円で、この平均値以上の世帯数と平均値以下の世帯数は半々でなく、全体の約55.8%の世帯が平均値よりも低くなっており、世帯数を二等分する中央値は455,560円となっていることと比較して、平成20年についてはどのようなことがわかりますか。」</p> <p>「では、現在、「不景気」とか「収入減」という言葉がよく社会問題として取り上げられているのですが、このことはどのようなデータをもとに調べたらよいでしょうか。」</p> <p>●【表1】を提示する。</p> <p>「これからデータの特徴としてどのようなことがわかりますか。」</p> <p>「平均値以外の代表値はよみとれますか。」</p> <p>「では、中央値はどのようになっていますか。」</p> <p>「平成16年と平成20年は同じ分布の傾向にあることがわかりましたが、他の年度はどうでしょうか。どのように比較すればよいでしょうか。」</p> <p>「ヒストグラムまたは度数分布多角形をかいて比較してみよう。」</p> <p>「これから何がわかりますか。また、代表値としては何をいいますか。」</p>	<p>「平成16年も平成20年も分布の傾向は変わらないのではないか。」</p> <p>「各年度、例えば、10年位遡り、毎年のデータをもとに分布の傾向を調べればよい。」</p> <p>「分布の傾向はほぼ似ているので、代表値で比較すればよい。」</p> <p>「年度毎の収入の推移を表したデータ。」</p> <p>「平均値が年々下がっている。」</p> <p>「中央値」</p> <p>「平均値と同様、年々下がっている。」</p> <p>「最頻値を求めて、傾向を調べる。」</p> <p>「ヒストグラムをかいて、傾向を調べる。」</p> <p>「度数分布多角形を用いて傾向を調べる。」</p> <p>「代表値で比較する。」</p> <p>「左に偏った分布なので、平均値で比較するよりも、中央値を代表値として、年度毎の推移を比較した方がよい。」</p> <p>「毎年の傾向として、「最頻値&lt;中央値&lt;平均値」の関係があるので、どの代表値を用いてもよい。」</p> <p>「表から、平均値と中央値はすぐによみとれるので、平均値または中央値で比較してよい。」</p>	<p>最頻値を提示する。</p> <p>表からヒストグラム、または度数分布多角形をかかせる。</p> <p>ヒストグラム、または度数分布多角形から「最頻値&lt;中央値&lt;平均値」の関係をつえる。</p>
<p>まとめ</p>	<p>「これは景気が良い時と不景気といわれる今日の家計の収入状況の特徴としてどのようなことがいえるでしょうか。」</p> <p>「収入の分布は左に偏った分布なので、収入の高い人は不況、収入減という状況は無関係である。」</p> <p>「平均値、中央値ともに年度毎に減少していることから、不景気と</p>	

	<p>いう影響は年度毎の比較からわかる。」</p> <p>「これから、収入の傾向は毎年、やや下の方に集中して分布している。これは景気が良い時も不景気といわれる今日も変わらない。」</p>
--	---

【提示資料】（総務省統計局の方から提示して頂いた資料）

		平均	I	II	III	IV	V	VI	VII	VIII	IX	X
平成 14 年	下限(万円)			197	280	349	415	494	583	694	839	1,068
	上限(万円)		197	280	349	415	494	583	694	839	1,068	
	平均値(万円)	587	138	239	315	381	453	535	638	762	939	1,466
平成 15 年	下限(万円)			194	278	348	412	489	576	680	809	1,023
	上限(万円)		194	278	348	412	489	576	680	809	1,023	
	平均値(万円)	570	136	238	314	379	448	530	625	741	905	1,380
平成 16 年	下限(万円)			196	279	340	405	480	566	673	808	1,026
	上限(万円)		196	279	340	405	480	566	673	808	1,026	
	平均値(万円)	563	136	239	310	372	442	520	616	738	906	1,351
平成 17 年	下限(万円)			195	273	336	401	473	556	655	793	1,004
	上限(万円)		195	273	336	401	473	556	655	793	1,004	
	平均値(万円)	554	136	235	304	368	436	513	603	720	886	1,335
平成 18 年	下限(万円)			190	265	329	392	460	545	648	783	997
	上限(万円)		190	265	329	392	460	545	648	783	997	
	平均値(万円)	551	132	229	298	360	424	501	596	712	879	1,381
平成 19 年	下限(万円)			188	268	332	396	461	540	643	778	997
	上限(万円)		188	268	332	396	461	540	643	778	997	
	平均値(万円)	553	130	228	301	363	426	499	589	706	875	1,413
平成 20 年	下限(万円)			190	269	332	393	461	546	641	774	991
	上限(万円)		190	269	332	393	461	546	641	774	991	
	平均値(万円)	547	135	231	301	362	426	502	592	704	872	1,348

【表 1】