

代表値と標準偏差を活用しよう

1. 統計学とは？

○世の中にあふれているデータ（数値）をまとめ、正しく分析する学問。

○統計学の3つの主な役割

①概念の（ ） ②（ ） ③仮説の（ ）や（ ）

○統計学の分野は、大きく分けると2つである。

・記述統計学 … 手元にあるデータを整理したり計算したりして分かりやすくする。

例：グラフや図で表す、平均値や標準偏差を計算する…など

※数学Iの「データの分析」で学習したことはこの分野

・推測統計学 … 手元にあるデータから、様々なものを予測する。

例：数百人のアンケート結果から全体の傾向を分析する、視聴率…など。

※詳しい理論は数学Bの第4章「確率分布と統計的な推測」にある。

2. 3つの代表値、どれを使う？

Q1：次の指標の英訳・意味・強みはどれでしょうか？

指標	英訳	意味	強み
平均値			
中央値			
最頻値			

【英訳】

A：average B：median C：mode

【意味】

ア：データにおいて最も個数が多い値

イ：データの値の総和をデータの個数で割った値

ウ：データを値の大きさ順に並べたとき、真ん中にくる値

【強み】

あ：極端な値に左右されることなく、「普通」がどれくらい分かる。

い：すべてのデータの値を使い、偏りをならすことで、「普通」がどれくらい分かる。

う：「人気」が何かが分かる。データが数値である必要がなく、言葉についても調べることができる。

Q2：データから次のことを知りたいとき、どの指標を考えればよいでしょう？

(1) 古川黎明の生徒の身長がどのくらいなのか知りたい。()

(2) 靴を売るときに、どのサイズの靴を仕入れるのがよいか知りたい。()

(3) 日本の社会人の年収がどの程度なのか知りたい。()

(4) 野球少年9人と社会人ピッチャー1人の球速の真ん中がどのくらいか知りたい。()

(5) 40万人が受験した100点満点の試験で全体的にとった点数がどのくらいか知りたい。()

(6) 交通事故の対策を考えるために、事故の主な原因が何かを知りたい。()

3. 標準偏差を使えば「スゴい度」が分かる！

○標準偏差 (Standard Deviation) … $\sqrt{\text{分散}}$ のこと。よく ([]) と表す。

※分散 = (平均との差=偏差)²の総和÷データの個数

→標準偏差を見ると、データの () の度合いが分かる。

(例) 県内にあるレストラン6店の1日の平均売上が20万円だったとする。

→A店では1日の売上が30万円であったとき、A店は優秀な店といえるか？

事例	A店	B店	C店	D店	E店	F店	総売上	平均売上	標準偏差
①	30						120	20	
②	30						120	20	

(単位は万円)

→「平均値より〇〇円だけ良い」だけで判断するのは危険だ、と分かる。

⇒「平均値」と「標準偏差」を両方考慮した数値で「スゴい度」を判断したい！

そこで、() を () で割った式、つまり $\frac{(\quad) - (\quad)}{(\quad)}$ を考えると、

○事例①の場合、式は $\frac{(\quad) - (\quad)}{(\quad)} \div (\quad) \leftarrow \frac{\text{A店の売上} - \text{平均売上}}{\text{標準偏差}}$ を計算する

○事例②の場合、式は $\frac{(\quad) - (\quad)}{(\quad)} \div (\quad)$

この式は、平均から標準偏差〇個分離れているかを表すので、①の場合は ()、

②の場合は () のように表す。

この「スゴい度」を () といい、標準得点でデータを考えることを () という。

Q3：国語と英語の2教科で100点満点のテストをしたとき、国語の平均は70点で標準偏差は12点、英語の平均は80点で標準偏差は4点だった。転校生のXくんは同じテストを受けて両方90点をとったとき、国語と英語のどっちがスゴいといえるだろうか？

(おまけ) 今計算した「標準得点」を使って、「標準得点×10+50」とした数値をいわゆる「偏差値」といいます。

上の(例)事例①と②それぞれのA店の売上の偏差値や、Q3のXくんの国語と英語それぞれの偏差値を求めてみよう。

代表値と標準偏差を活用しよう

1. 統計学とは？

○世の中にあふれているデータ（数値）をまとめ、正しく分析する学問。

○統計学の3つの主な役割

①概念の（計量化） ②（予測） ③仮説の（検証）や（分類）

○統計学の分野は、大きく分けると2つである。

・記述統計学 … 手元にあるデータを整理したり計算したりして分かりやすくする。

例：グラフや図で表す、平均値や標準偏差を計算する…など

※数学Iの「データの分析」で学習したことはこの分野

・推測統計学 … 手元にあるデータから、様々なものを予測する。

例：数百人のアンケート結果から全体の傾向を分析する、視聴率…など。

※詳しい理論は数学Bの第4章「確率分布と統計的な推測」にある。

2. 3つの代表値、どれを使う？

Q1：次の指標の英訳・意味・強みはどれでしょうか？

指標	英訳	意味	強み
平均値	A	イ	い
中央値	B	ウ	あ
最頻値	C	ア	う

【英訳】

A：average B：median C：mode

【意味】

ア：データにおいて最も個数が多い値

イ：データの値の総和をデータの個数で割った値

ウ：データを値の大きさ順に並べたとき、真ん中にくる値

【強み】

あ：極端な値に左右されることなく、「普通」がどれくらい分かる。

い：すべてのデータの値を使い、偏りをなすことで、「普通」がどれくらい分かる。

う：「人気」が何かが分かる。データが数値である必要がなく、言葉についても調べることができる。

Q2：データから次のことを知りたいとき、どの指標を考えればよいでしょう？

(1) 古川黎明の生徒の身長がどのくらいなのか知りたい。（平均値）

(2) 靴を売るときに、どのサイズの靴を仕入れるのがよいか知りたい。（最頻値）

(3) 日本の社会人の年収がどの程度なのか知りたい。（中央値）

(4) 野球少年9人と社会人ピッチャー1人の球速の真ん中がどのくらいか知りたい。（中央値）

(5) 40万人が受験した100点満点の試験で全体的にとった点数がどのくらいか知りたい。（平均値）

(6) 交通事故の対策を考えるために、事故の主な原因が何かを知りたい。（最頻値）

3. 標準偏差を使えば「スゴい度」が分かる！

○標準偏差（Standard Deviation）… $\sqrt{\text{分散}}$ のこと。よく（ σ [シグマ]）と表す。

※分散 = (平均との差=偏差)²の総和÷データの個数

→標準偏差を見ると、データの（バラツキ）の度合いが分かる。

(例) 県内にあるレストラン6店の1日の平均売上が20万円だったとする。

→A店では1日の売上が30万円であったとき、A店は優秀な店といえるか？

事例	A店	B店	C店	D店	E店	F店	総売上	平均売上	標準偏差
①	30	10	10	20	20	30	120	20	8.16
②	30	5	5	5	35	40	120	20	15.28

(単位は万円)

→「平均値より〇〇円だけ良い」だけで判断するのは危険だ、と分かる。

⇒「平均値」と「標準偏差」を両方考慮した数値で「スゴい度」を判断したい！

そこで、（平均からのズレ）を（標準偏差）で割った式、つまり $\frac{\text{（平均からのズレ）}}{\text{（標準偏差）}}$ を考えると、

○事例①の場合、式は $\frac{(30) - (20)}{(8.16)} \doteq (+1.23)$ ← $\frac{\text{A店の売上} - \text{平均売上}}{\text{標準偏差}}$ を計算する

○事例②の場合、式は $\frac{(30) - (20)}{(15.28)} \doteq (+0.65)$

この式は、平均から標準偏差〇個分離れているかを表すので、①の場合は（+1.23SD）、

②の場合は（+0.65SD）のように表す。

この「スゴい度」を（標準得点）といい、標準得点でデータを考えることを（標準化）という。

Q3：国語と英語の2教科で100点満点のテストをしたとき、国語の平均は70点で標準偏差は12点、英語の平均は80点で標準偏差は4点だった。転校生のXくんは同じテストを受けて両方90点をとったとき、国語と英語のどっちがスゴいといえるだろうか？

(おまけ) 今計算した「標準得点」を使って、「標準得点×10+50」とした数値をいわゆる「偏差値」といいます。

上の(例)事例①と②それぞれのA店の売上の偏差値や、Q3のXくんの国語と英語それぞれの偏差値を求めてみよう。