

Statistical Room



平均値と中央値

今回は平均値と中央値についてのお話です。

平均値とは、データの合計をデータの個数で割って得られる値

中央値とは、データを大きさの順に並べ替えたとき、ちょうど順番が真ん中になる値です。

具体例を見てみましょう。

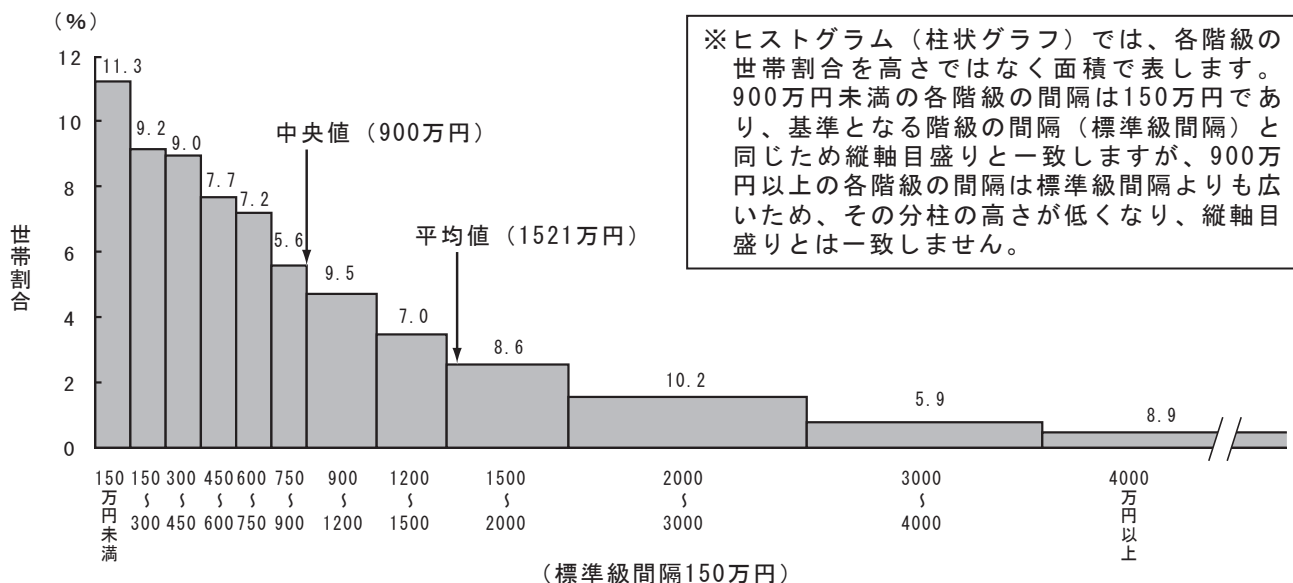
5人の小学生がいて、あるテストの点数が30点、40点、50点、60点、70点だったとき、平均値は「 $(30+40+50+60+70) \div 5 = 50$ 点」となります。一方、中央値は今回はデータが5つありますので、真ん中（3つ目）の値である50点となります。このケースは平均値と中央値が同じになりました。

では、次の場合はどうでしょう。

同じく5人の小学生のテストの点数が30点、35点、40点、40点、90点だったとき、平均値は「 $(30+35+40+40+90) \div 5 = 47$ 点」となり、90点以外の4人は平均値以下となります。中央値は、真ん中の値ですので40点となります。

このように、平均値の場合は他の値と比べて極端に高い（もしくは低い）値があることによって、影響を受けてしまいます。中央値の場合は、真ん中の値ですので、そのような影響は受けません。

では、最後に総務省統計局が実施している平成21年全国消費実態調査の二人以上の世帯の貯蓄額を見てみましょう。



このように偏りがあるデータの場合、平均値は当然、高額貯蓄世帯の影響により高い方に引っ張られますので、平均値だけではなく併せて中央値も見るべきです。

平均値と中央値、それぞれの特徴を知った上で使用することが重要です。