

# Statistical Room



## 標準偏差

データを見るときに、平均値は一般的によく使われますが、平均値の他に、「標準偏差」があり、データのバラツキが分かります。

具体例を挙げて考えていきましょう。自宅から目的地まで行くのに、「バスのみで行く方法」と「電車と徒歩で行く方法」があり、所要時間が下表のようになった場合、それぞれについて平均時間を計算すると、ともに32分と同じになります。

しかし、平均時間は同じでも所要時間のバラツキは異なります。そこで、そのバラツキを数値で表す代表的な方法として「標準偏差」が登場します。

	所要時間		所要時間 - 平均		(所要時間 - 平均) <sup>2</sup>	
	バス	電車+徒歩	バス	電車+徒歩	バス	電車+徒歩
1回目	34	31	34-32 = 2	-1	(2) <sup>2</sup> = 4	1
2回目	27	31	27-32 = -5	-1	(-5) <sup>2</sup> = 25	1
3回目	44	32	12	0	144	0
4回目	21	35	-11	3	121	9
5回目	33	24	1	-8	1	64
6回目	39	33	7	1	49	1
7回目	26	36	-6	4	36	16
8回目	34	29	2	-3	4	9
9回目	38	40	6	8	36	64
10回目	24	29	-8	-3	64	9
合計	320	320	0	0	484	174
合計÷データ個数	32.0	32.0	0.0	0.0	48.4	17.4
√合計÷データ個					7.0	4.2

平均

④標準偏差

上表のとおり、①各データと平均の差を計算し、②それを2乗した上で合計し、③その合計をデータ個数で割った後に④平方根をとると標準偏差になります。

標準偏差が小さいと各データは、平均値付近に固まっています。標準偏差が大きいと、各データは平均値から離れたところに散らばっています。つまり、今回の場合は、「バス」と「電車+徒歩」を比べると「電車+徒歩」の方が、目的地までの所要時間が安定しているということになります。

安定して時間通りに着きたい場合は「電車+徒歩」を利用するのが良いでしょう。

最後にヒストグラムにしてみます。数値を見るだけでなく、グラフにしてどのような分布になっているかを確認することも大切です。

