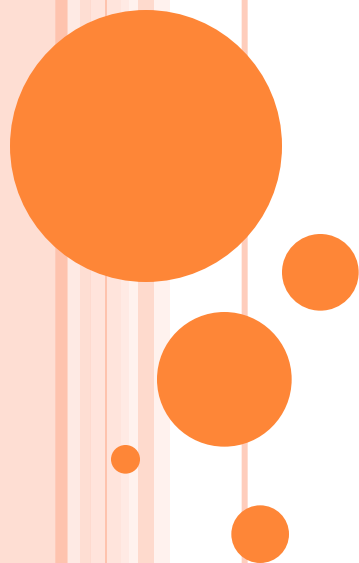


統計学

第2回

西山



第1回の復習

1. 平均値が100、標準偏差が20である分布の形を大雑把に描きなさい。
2. 分布図の右端(=最大値)は、いくら位であるか、適切な値を書き加えておきなさい。また、最小値についても同様に書き加えなさい。



第1回の復習

- (0) 5個のデータ、1, 2, 3, 4, 5の標準偏差を求めなさい。
- (1) すべてのデータに一定の数値(たとえば10、-10)を加えた場合、平均値はどのように変化するか？
- (2) すべてのデータに一定の数値(たとえば2、0.5)をかけた場合、平均値はどうなるか？
- (3) すべてのデータに一定の数値(たとえば10、-10)を加えた場合、標準偏差はどのように変化するか？
- (4) すべてのデータに一定の数値(たとえば2、0.5)をかけた場合、標準偏差はどうなるか？

ここまでできたら次の問題

- (5) 偏差の合計は常にゼロである。
- (6) 分散 = 二乗の平均 - 平均の二乗

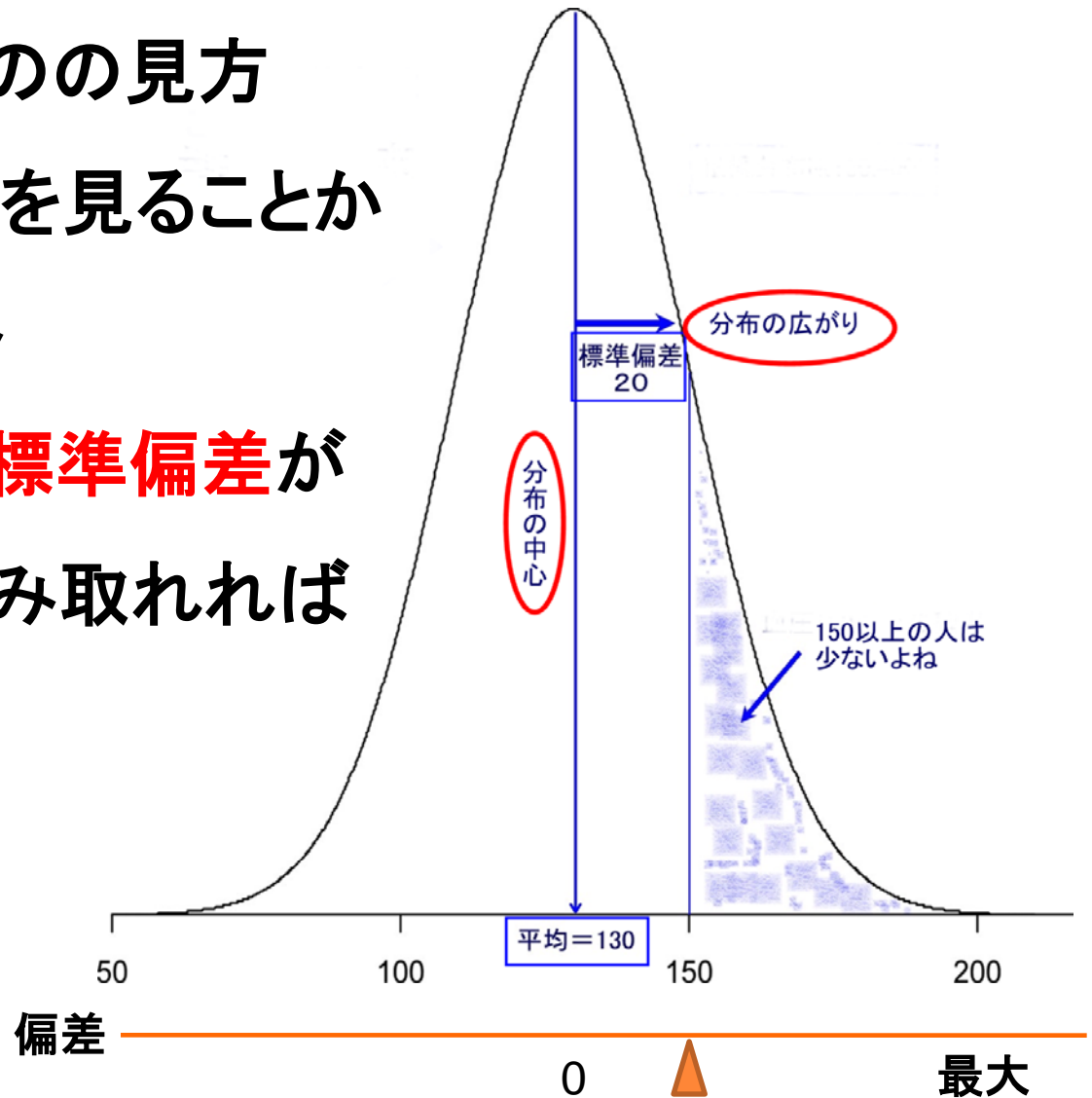
データの公式
と呼んでいます



第1回のまとめ

統計的なものの見方

- 分布の形を見ることからスタート
- 平均値と標準偏差が図から読み取ればよし



第2回目の目標

1. 確率的な考え方
2. データの平均と確率分布の平均
3. データの標準偏差と確率分布のそれ

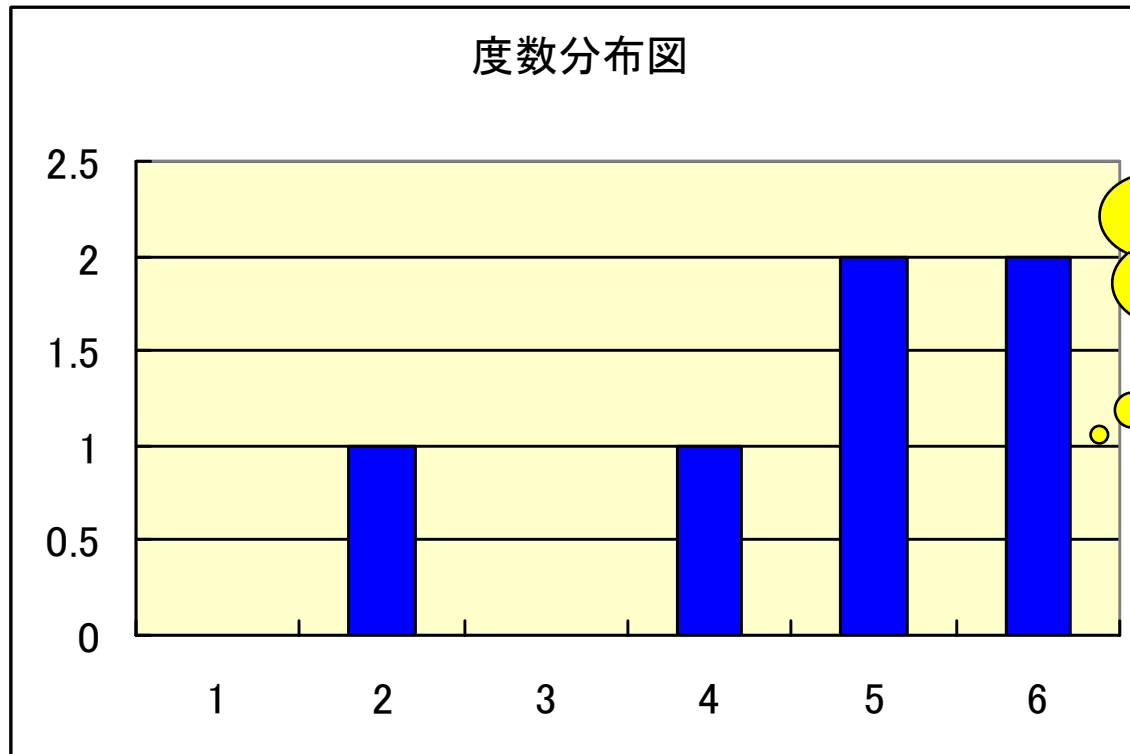
教科書： 第2章の頁49～53

第2章の頁67～71



データの分布

サイコロを6回振ると、 **6、4、2、6、5、5**



目の数の平均
と標準偏差、
出せますね

$$\bar{X} = 4.67$$

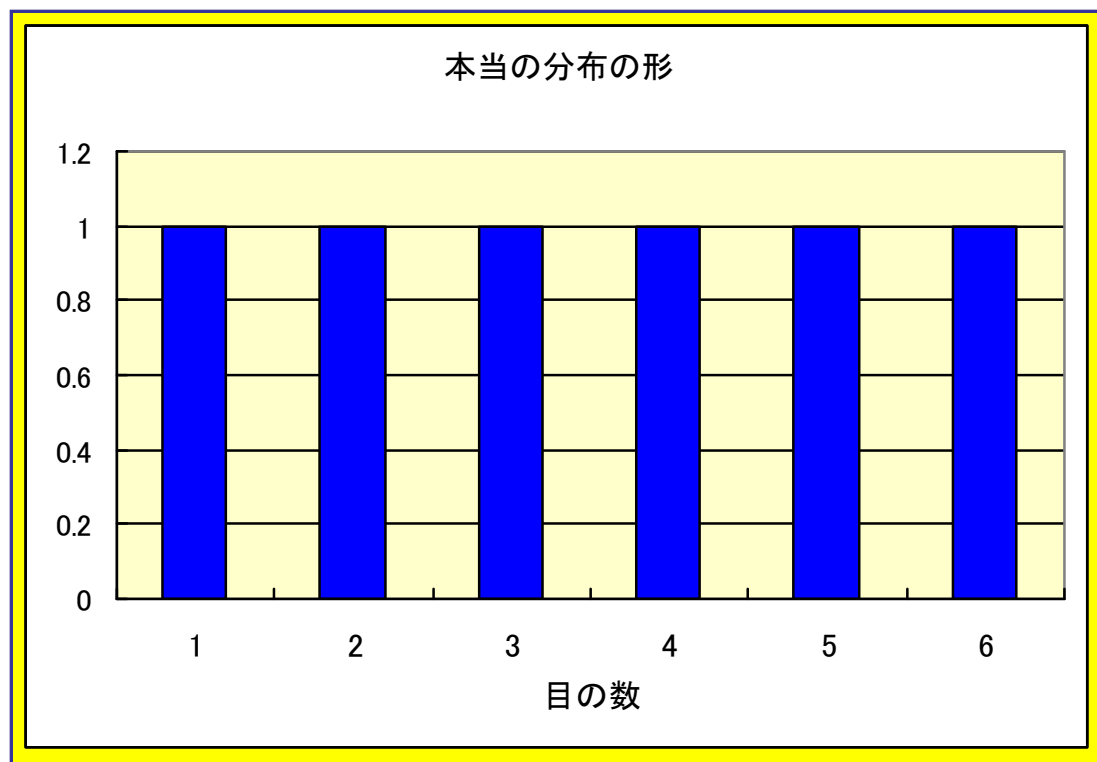
$$S^2 = 1.89$$

$$S = 1.37$$

1の目と3の目が出る確率はゼロである！？

正しいサイコロの目の出方

確率分布と呼びます



理屈の割合のことを確率といって

$P(X=1)=1/6$
と書きます。

目の数をXとすると割合の決まった変数だから確率変数と呼んでいます。

縦軸の値は1回ずつというより1/6ずつ、ですね。

同じ分布ではあるが

データの分布 vs 理論の分布



ヒストグラム



確率分布



観察事実

理論的な分布

分布の中心と広がり

平均と標準偏差



確率分布の平均 期待値といいます

『正常なコインを投げて表だと君の勝ちだ。私が千円あげる。だが裏だと千円もらおう！』



どちらが得をするか？
何度も反復すると

割合 × 値の合計..
これは平均計算です
教科書68－69ページ
を見よ.

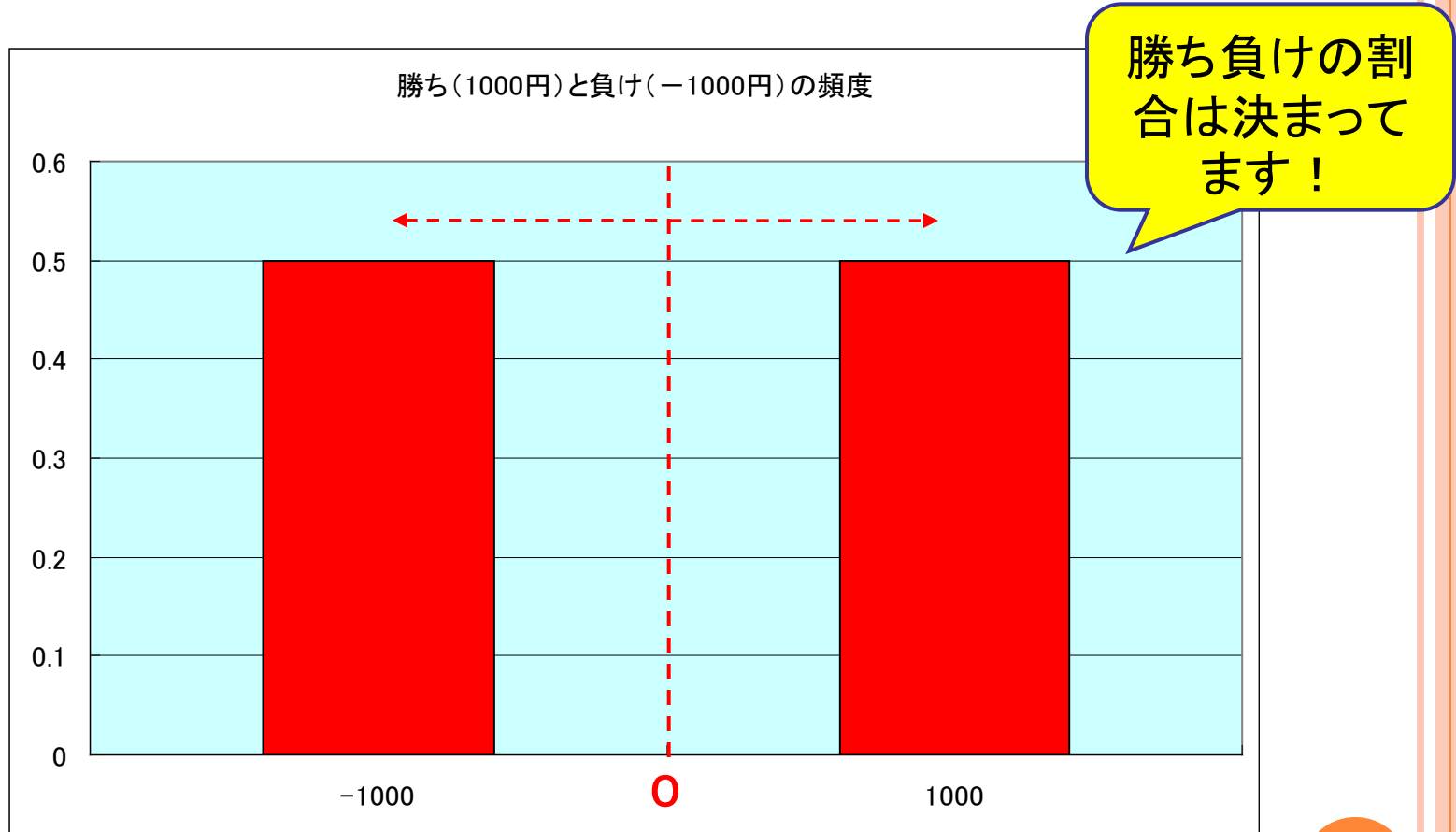
$$E[\text{利益}] = 0.5 \times (+1000) + 0.5 \times (-1000) = 0$$

勝ち負け半々だから、損も得もしない理屈.

数学的期待値といいます。意味は「理屈の平均」です。



勝ち負けの確率分布図



回数でなく確率を示しているから**確率分布**です

実際に10回やってみると

表が7回、裏が3回出た。合計で4000円もうけた。平均ではどれだけトクをしたか？

$$\begin{aligned}\bar{X} &= \frac{1000 + 1000 + (-1000) + \cdots + 1000}{10} \\ &= \frac{7 \times 1000 + 3 \times (-1000)}{10} && \text{合計} \div \text{回数} \\ &= \frac{7}{10} \times 1000 + \frac{3}{10} \times (-1000) && \text{値} \times \text{割合} \\ &= 400\end{aligned}$$

理屈と
違う！！

$$E[X] = 0.5 \times 1000 + 0.5 \times (-1000) = 0 \quad \leftarrow \text{割合} \times \text{値の合計}$$

Xの平均、あるいはXの**期待値**と呼びます。

確率分布の散布度

『1000円から10000円に上げよう』



X	P
1000	0.5
-1000	0.5

X	P
10000	0.5
-10000	0.5



平均値は、双方とも $E[X]=0$



確率分布のばらつき

	利得	偏差	二乗偏差	確率
	-1000	-1000	1000000	0.5
	1000	1000	1000000	0.5
平均	0	0	1000000	
			↓	
		分散 =	1000000	
		標準偏差 =	1000	

利得をXと置くと

$$E[X] = 0$$

$$V[X] = 10^6$$

$$SD[X] = 1000$$

同じく...

$$E[X] = 0$$

$$V[X] = 10^8$$

$$SD[X] = 10000$$

	利得	偏差	二乗偏差	確率
	-10000	-10000	100000000	0.5
	10000	10000	100000000	0.5
平均	0	0	100000000	
			↓	
		分散 =	1E+08	
		標準偏差 =	10000	

教科書： 頁67～71