The Law of Large Numbers
And
Central Limit Theorem

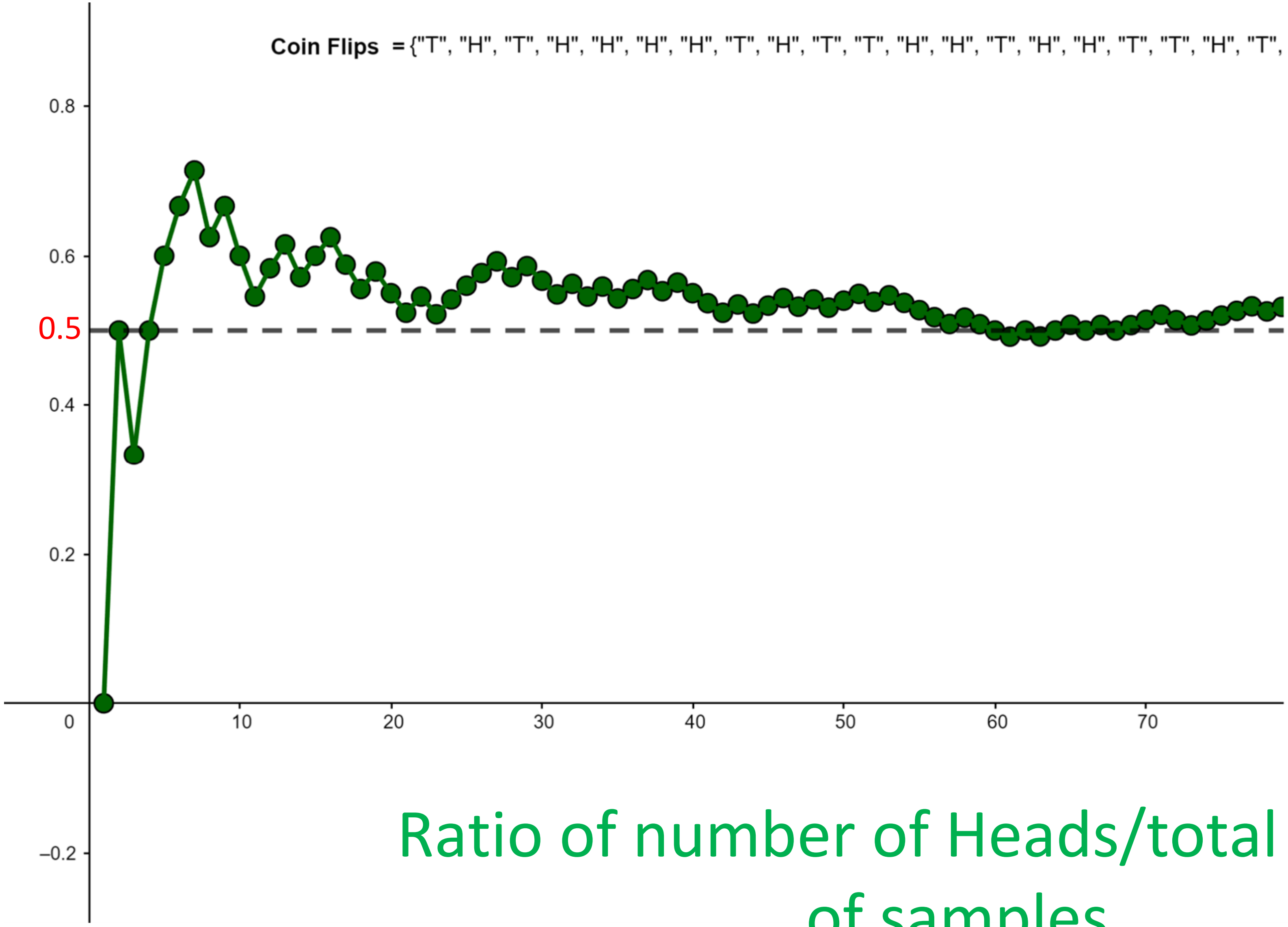
Dam Quang Tuan

Learning Goals

- Law of large numbers.
- Central limit theorem.
- Probabilities of averages and sums of independent identically-distributed random variables.

Law of Large Numbers

Example

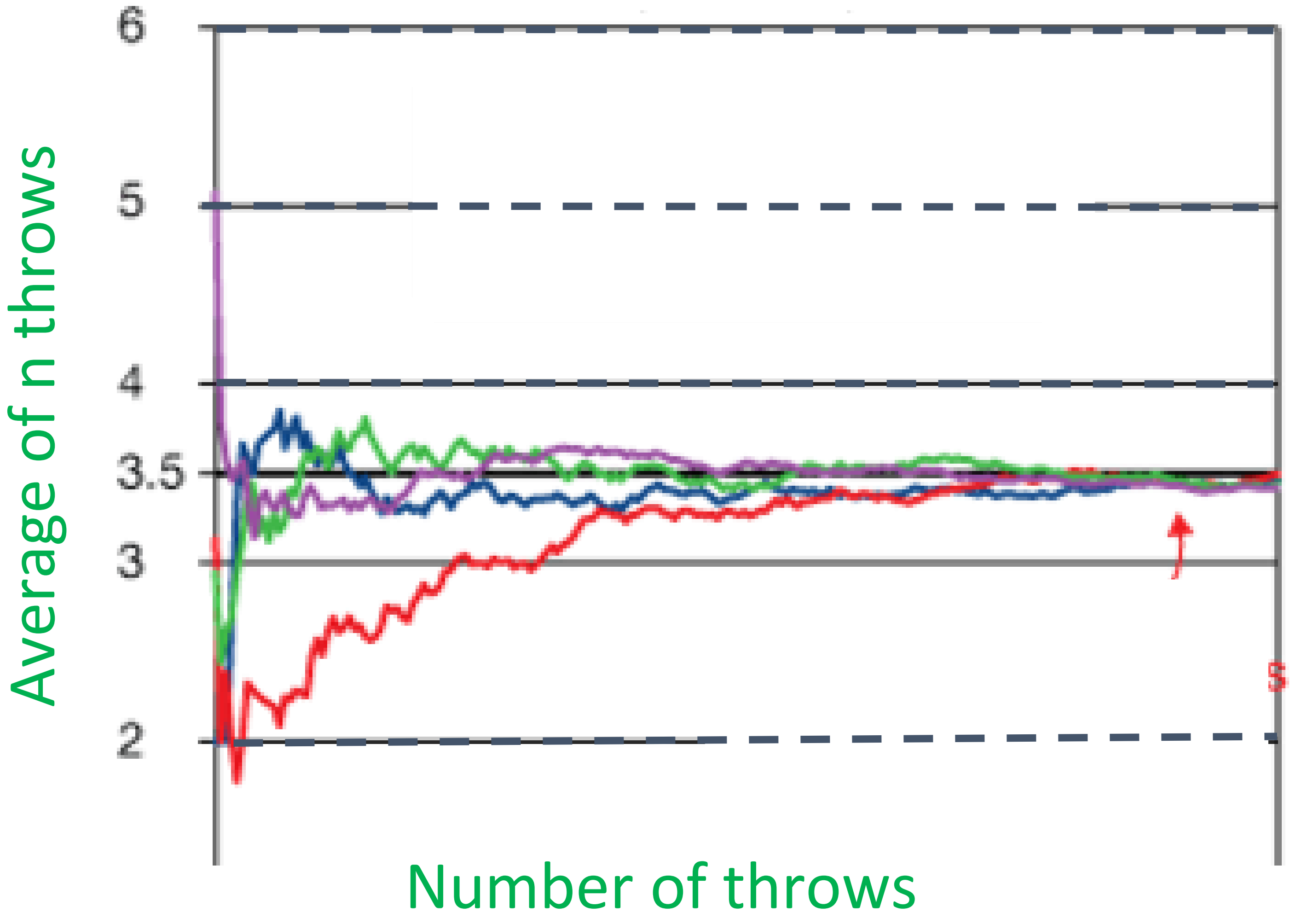


H: Head

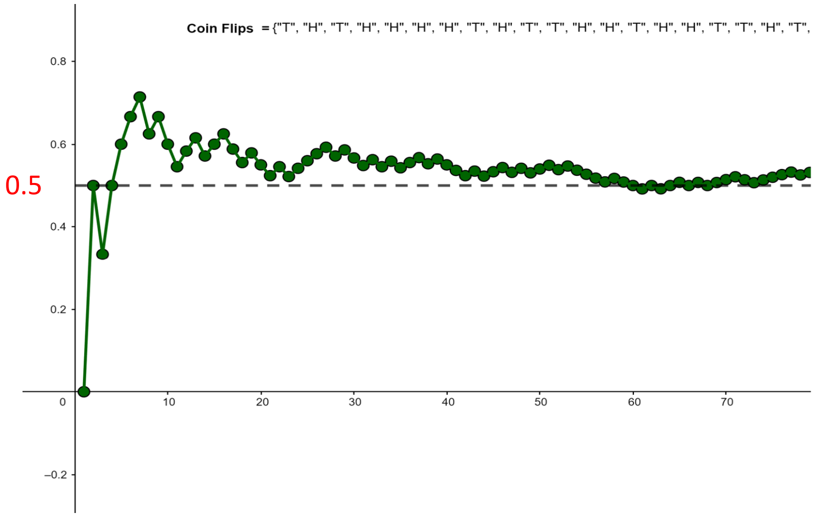
T: Tail

Ratio of number of Heads/total number of samples

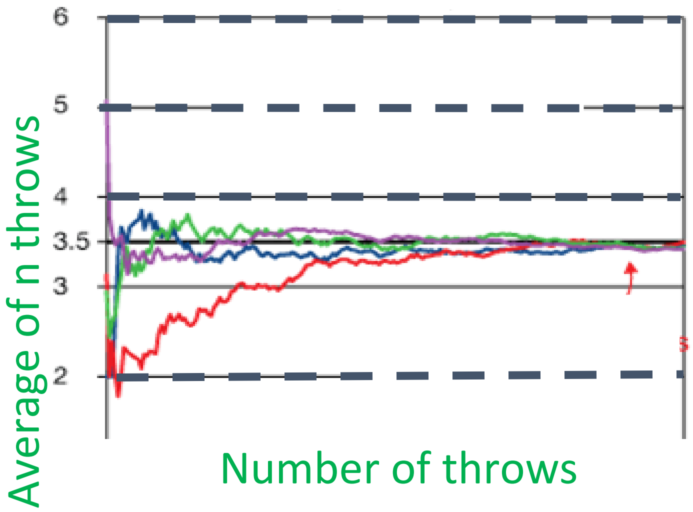
Example



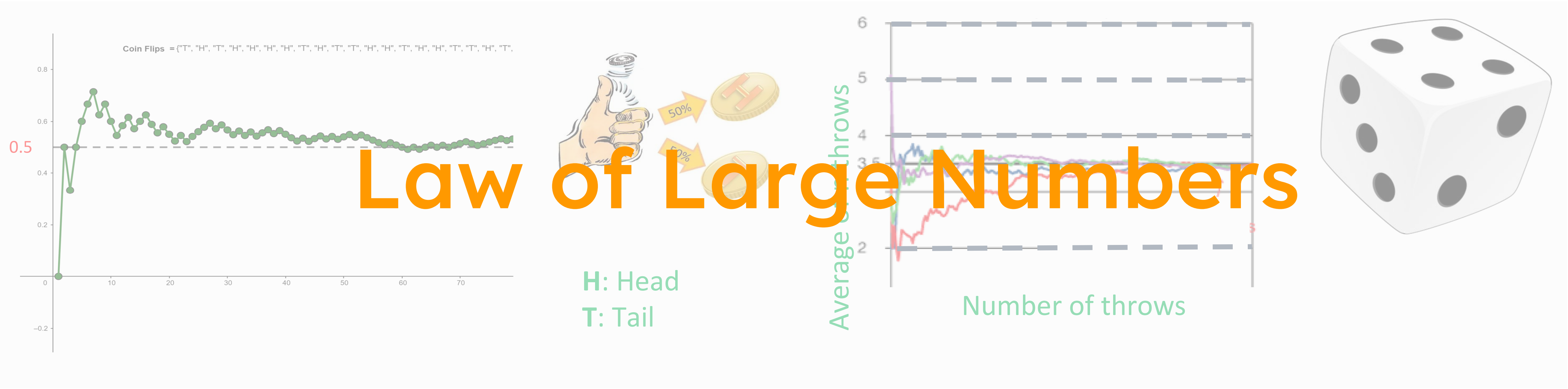
Example



H: Head
T: Tail



Example

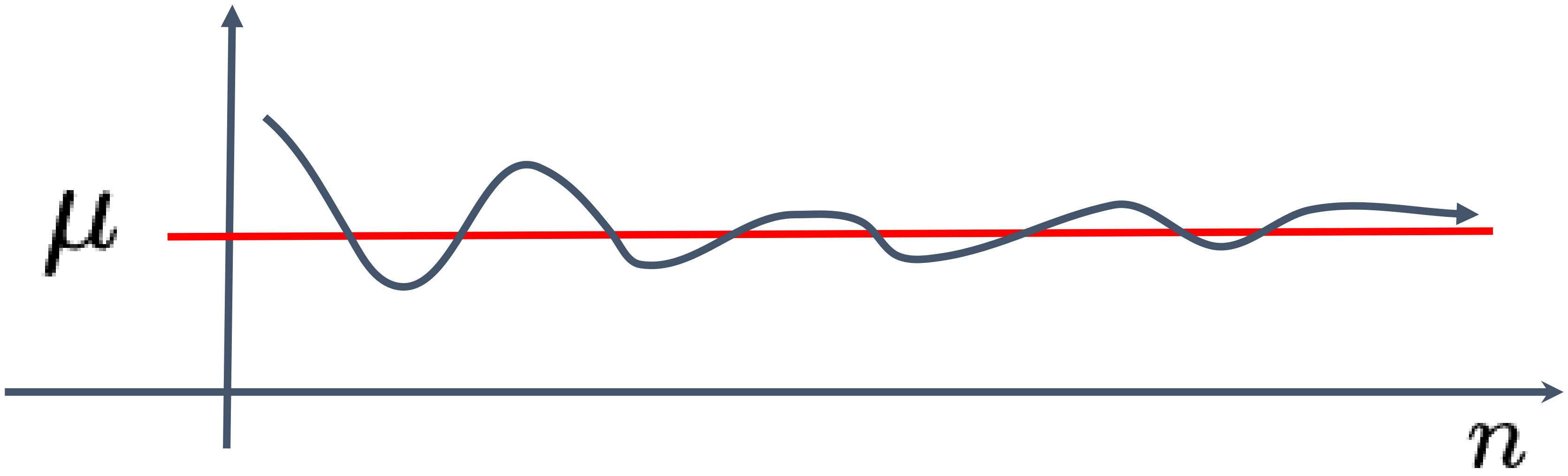


Law of Large Numbers

The Law of Large Numbers

Suppose $X_1, X_2, \ldots, X_n$ **are independent random variables with the same underlying distribution with mean** μ

$$\frac{X_1+X_2+\ldots+X_n}{n} \xrightarrow{n\to\infty} \mu$$

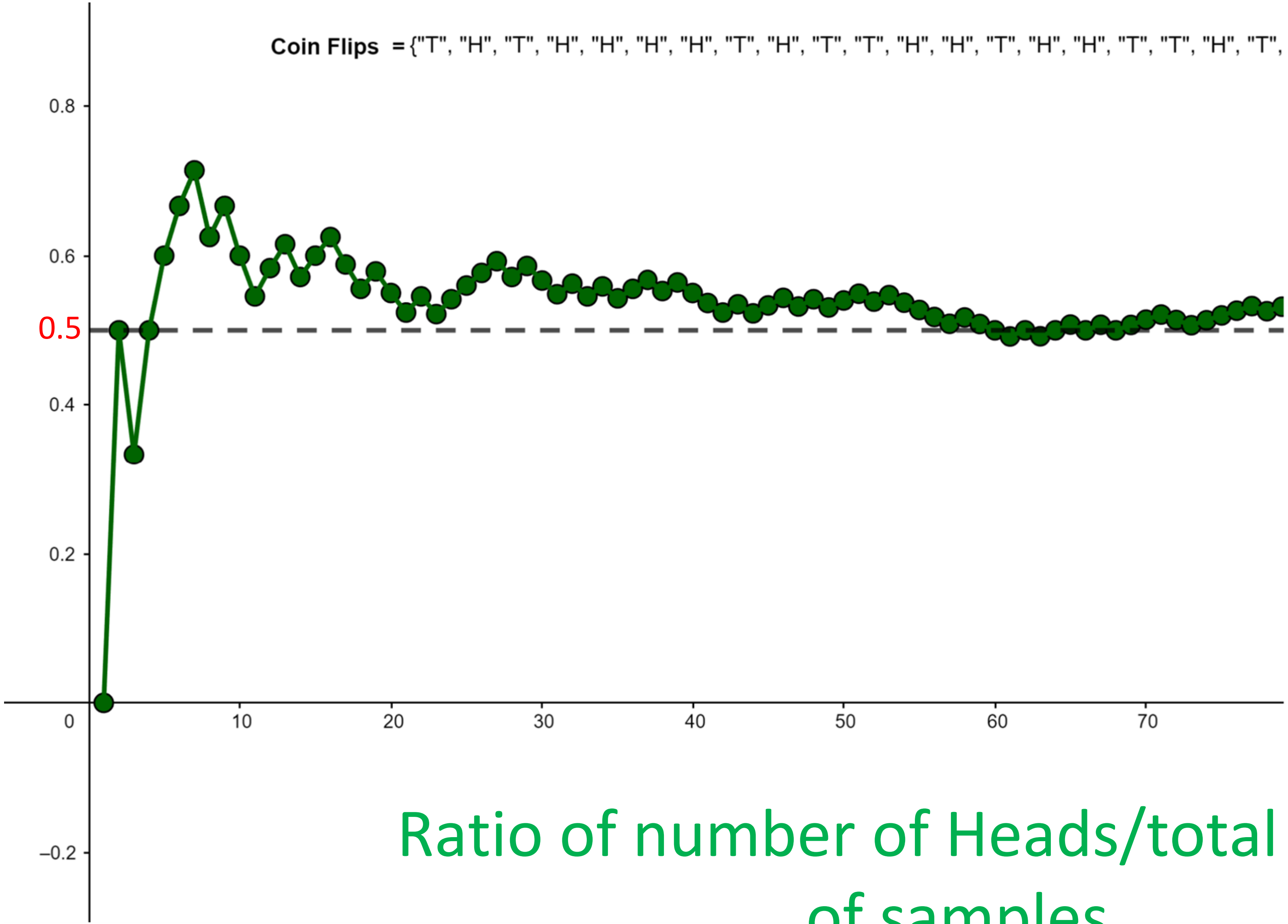


The Law of Large Numbers

Theorem (Law of Large Numbers): Suppose $X_1, X_2, \ldots, X_n, \ldots$ are i.i.d. random variables with mean μ. For each n, let $\overline{X}_n$ be the average of the first n variables. Then for any $a > 0$, we have

$$\lim_{n\to\infty} P(|\overline{X}_n - \mu| < a) = 1.$$

Example

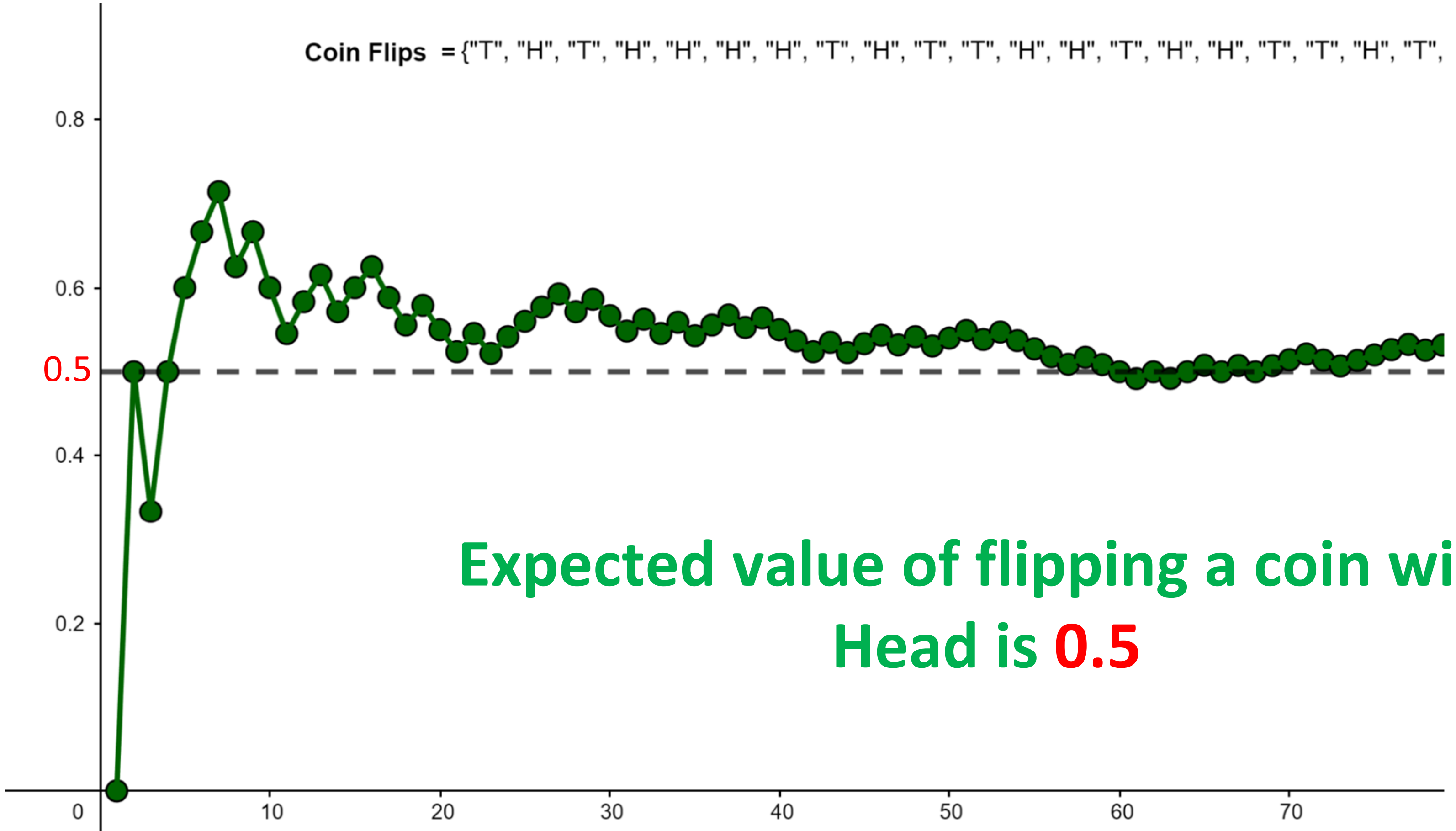


H: Head

T: Tail

Ratio of number of Heads/total number of samples

Example



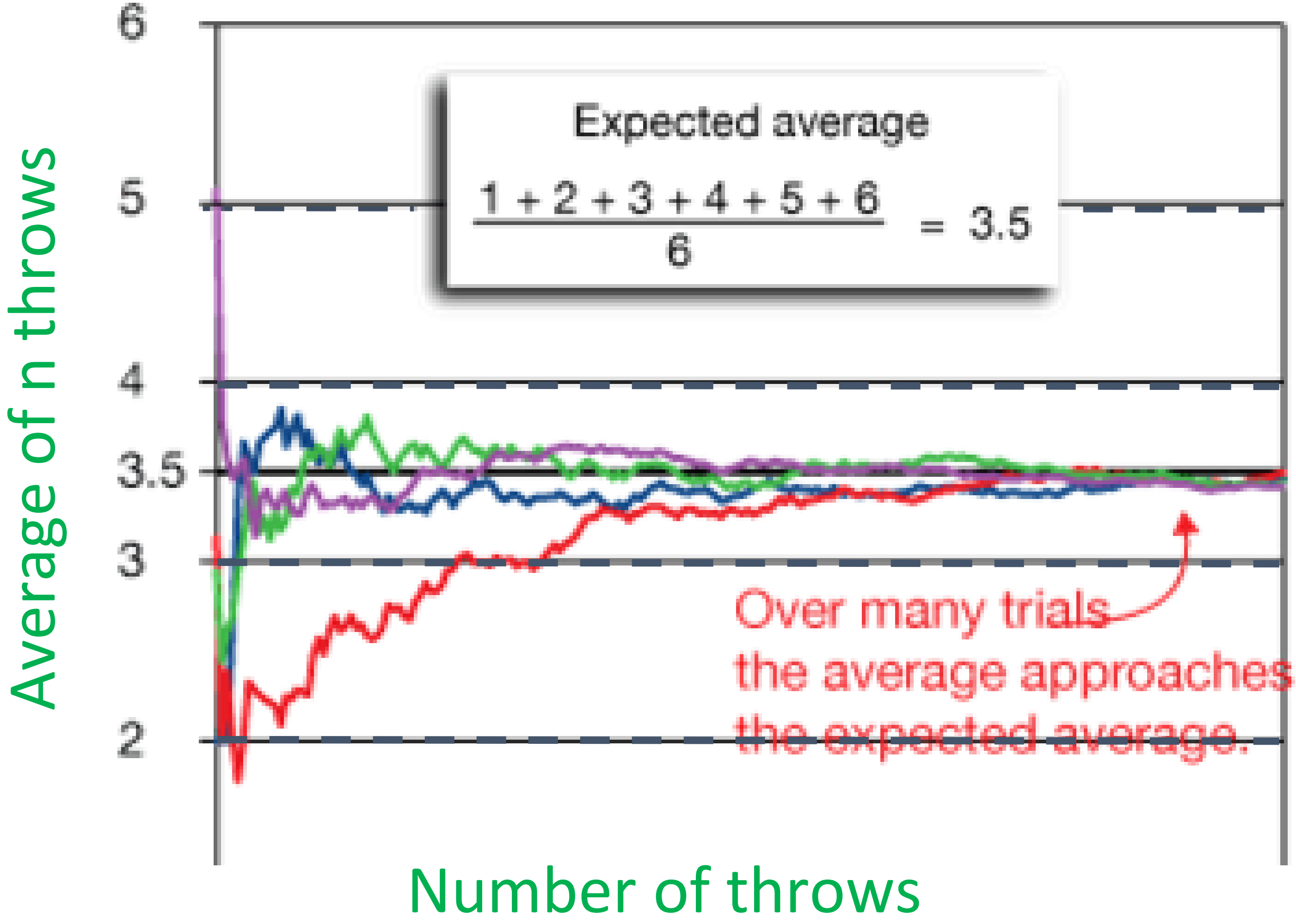
Expected value of flipping a coin with Head is 0.5

H: Head

T: Tail

$$X_i = \begin{cases} 1, & \text{if } H \\ 0, & \text{if } T \end{cases} \qquad \frac{X_1 + X_2 + \ldots + X_n}{n} \xrightarrow{n \to \infty} \mu$$

Example



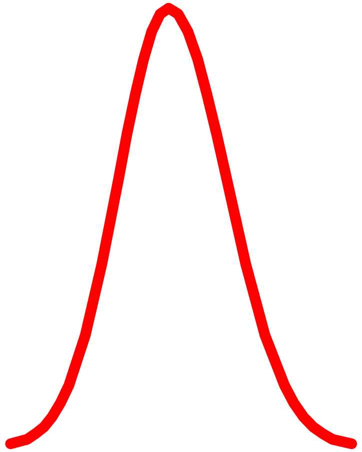
$$\frac{X_1+X_2+\ldots+X_n}{n} \xrightarrow{n\to\infty} \mu$$

Central Limit Theorem

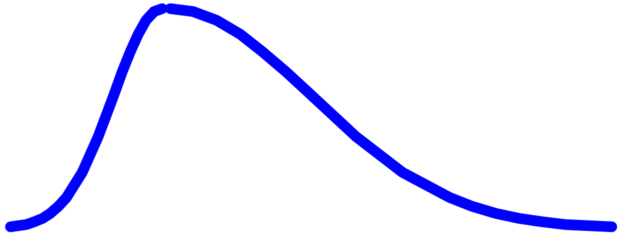
Example

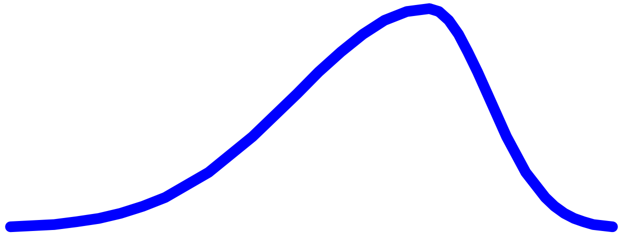
Galton Board

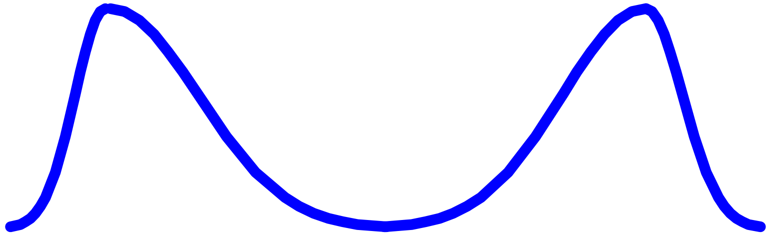
Example



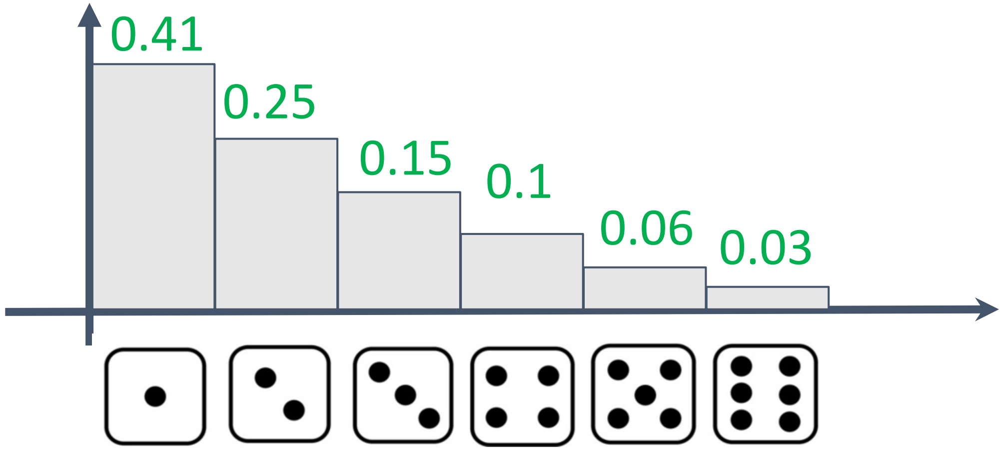
Gaussian distribution







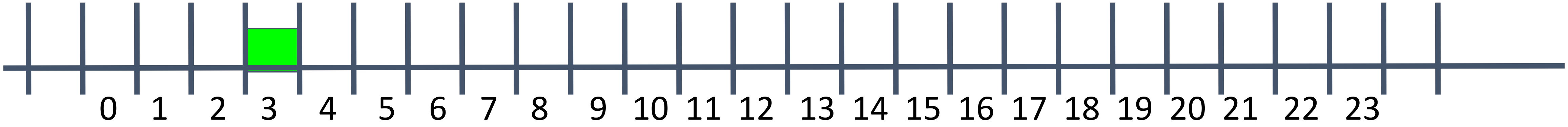
Example



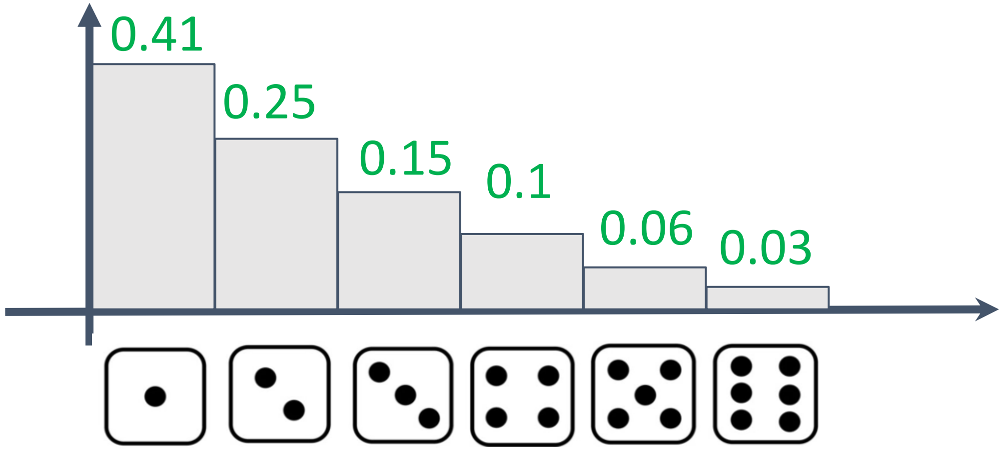
Sum=3

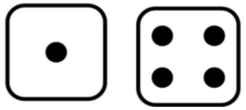
2 Dice

#Sum=1



Example

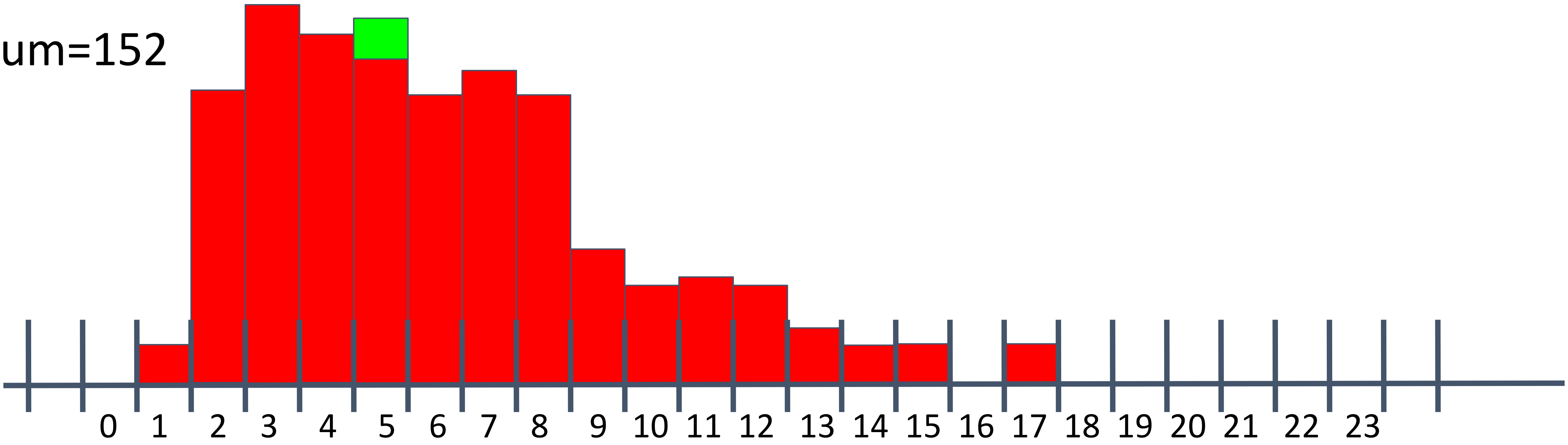




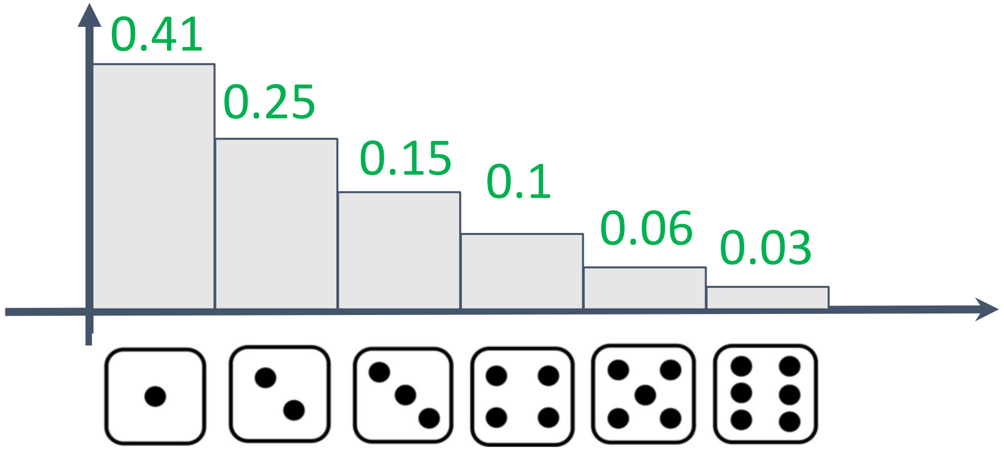
Sum=5

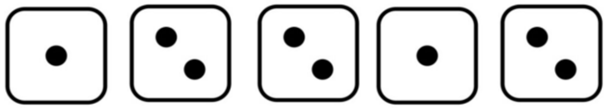
2 Dice

#Sum=152



Example

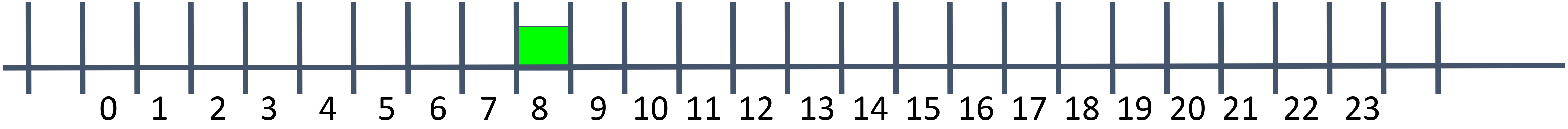




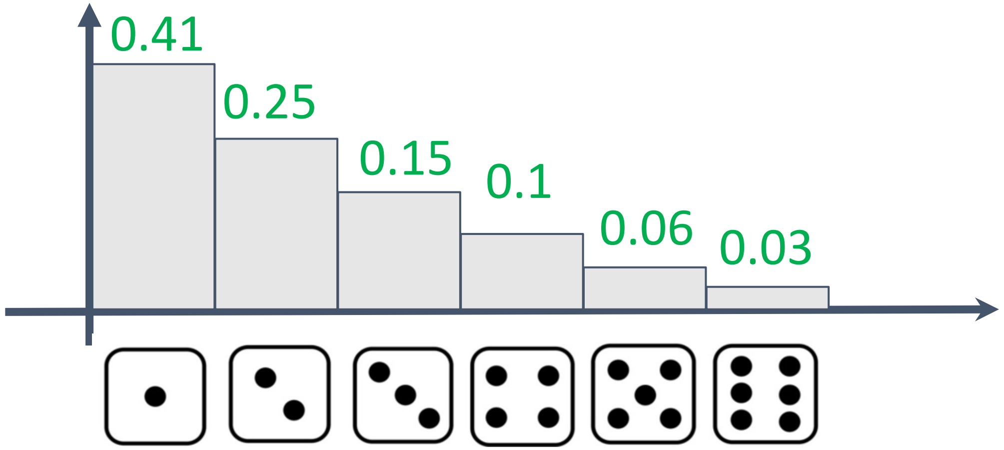
Sum=8

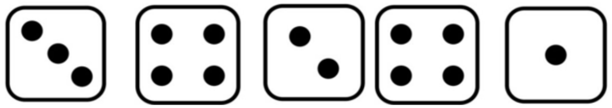
5 Dice

#Sum=1



Example

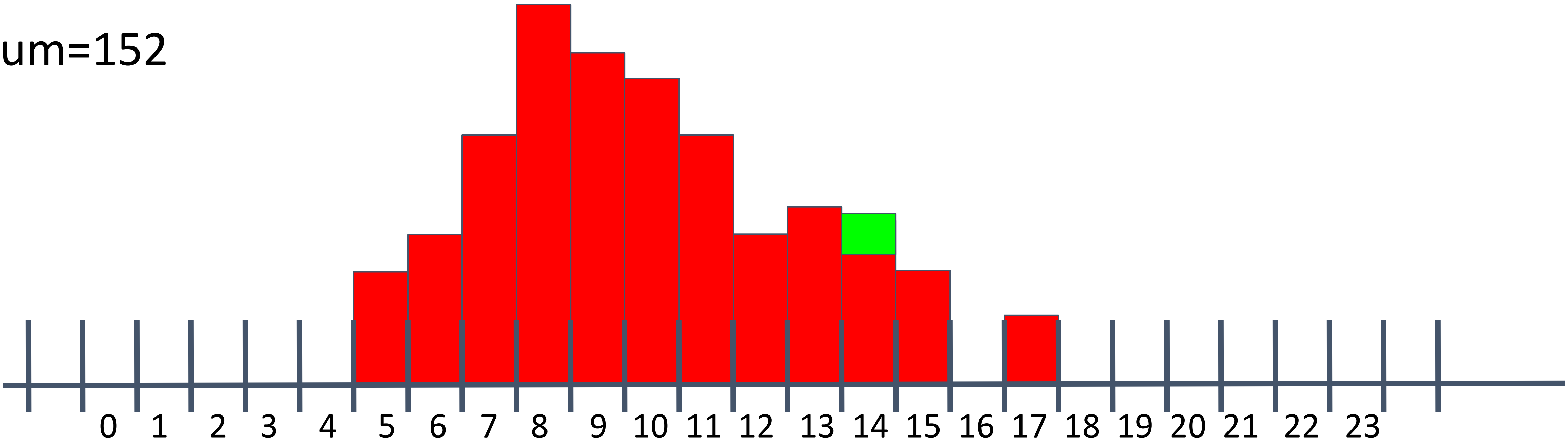




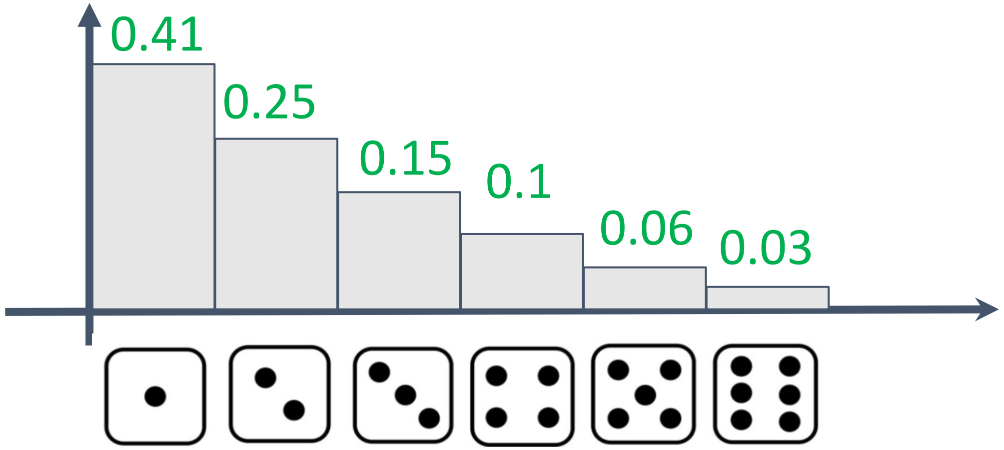
Sum=14

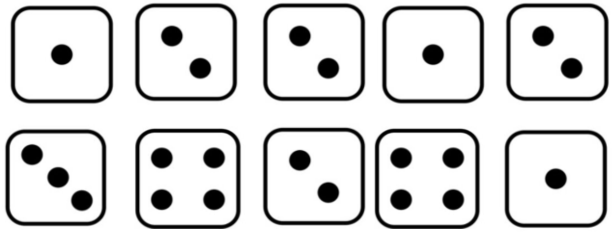
5 Dice

#Sum=152



Example

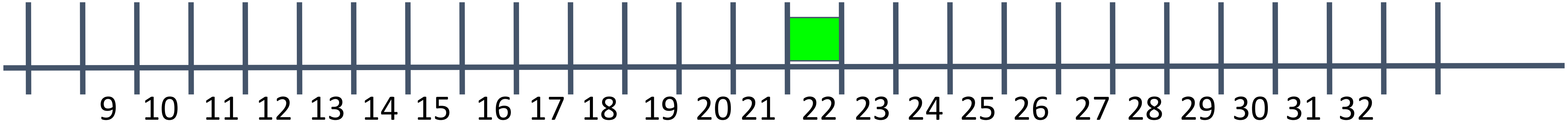




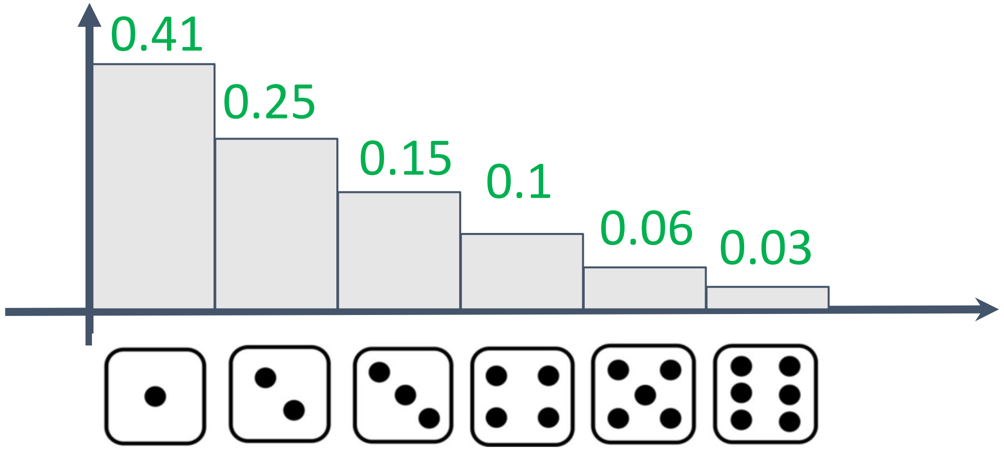
Sum=22

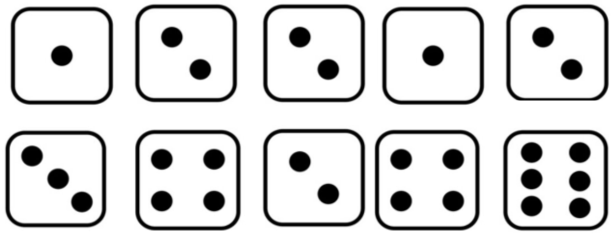
10 Dice

#Sum=152



Example

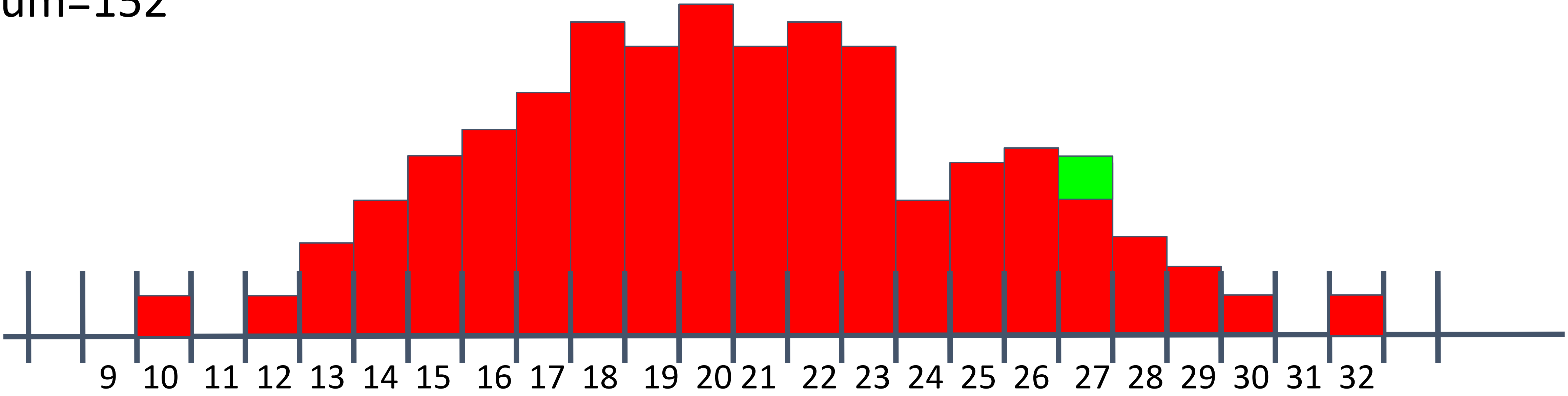




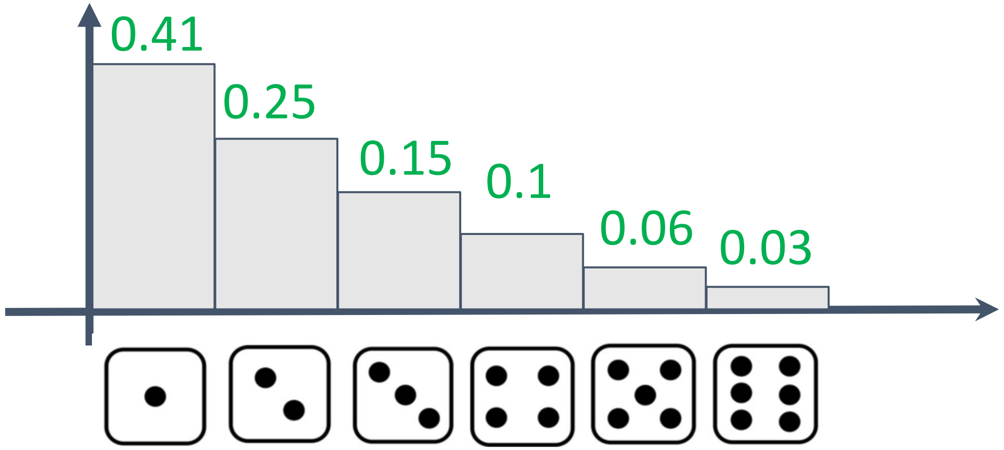
Sum=27

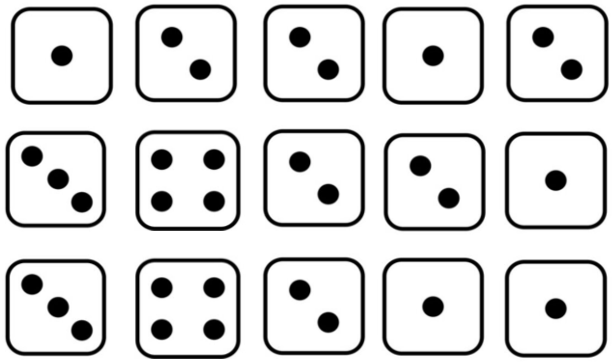
10 Dice

#Sum=152



Example

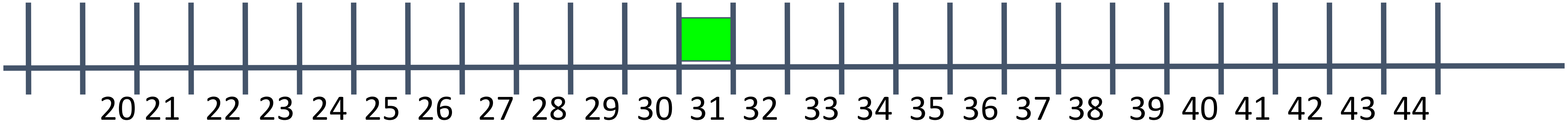




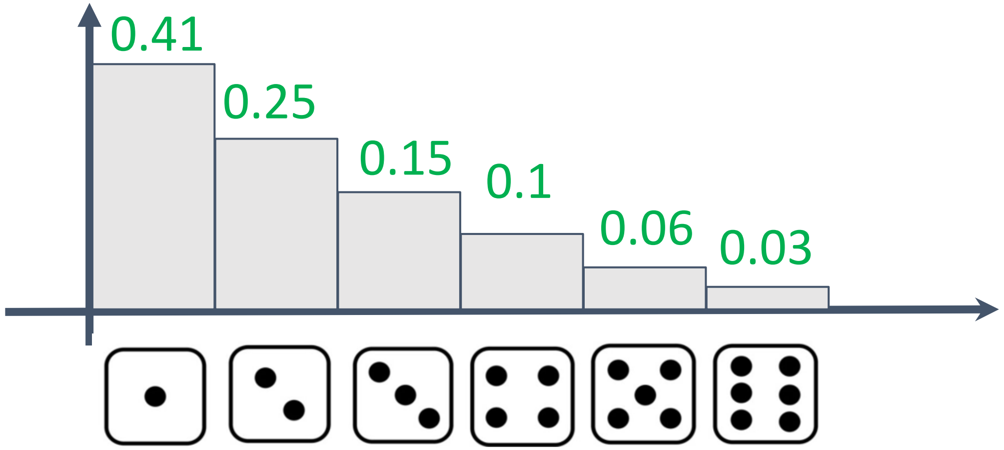
Sum=31

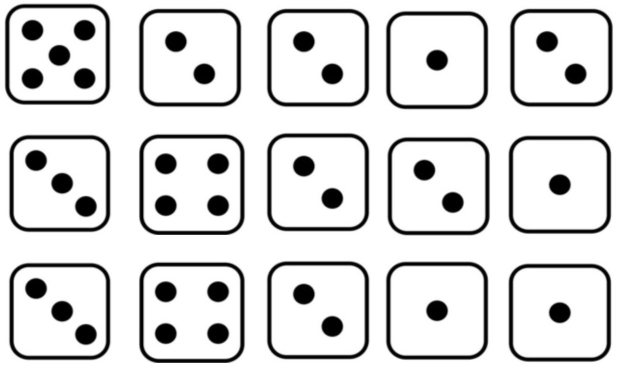
#Sum=1

15 Dice



Example

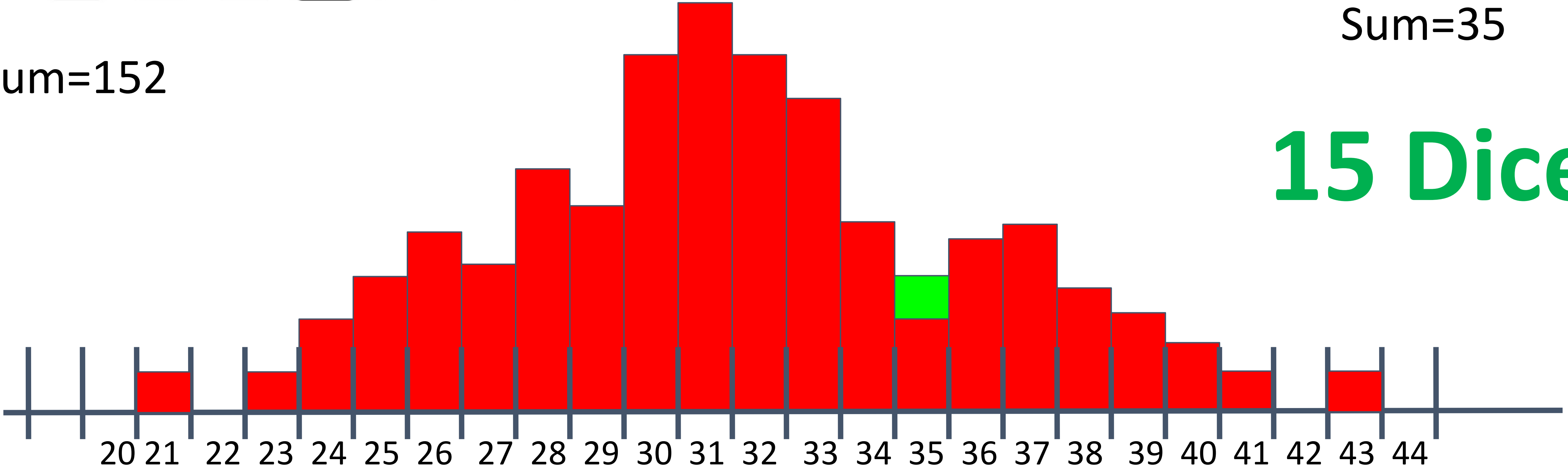




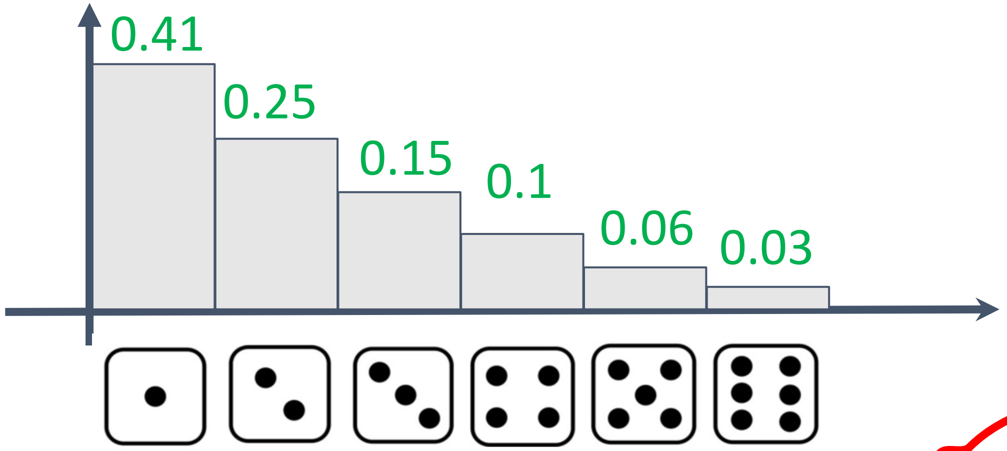
Sum=35

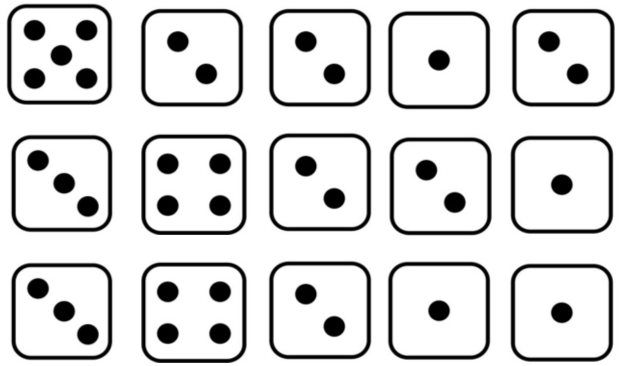
#Sum=152

15 Dice



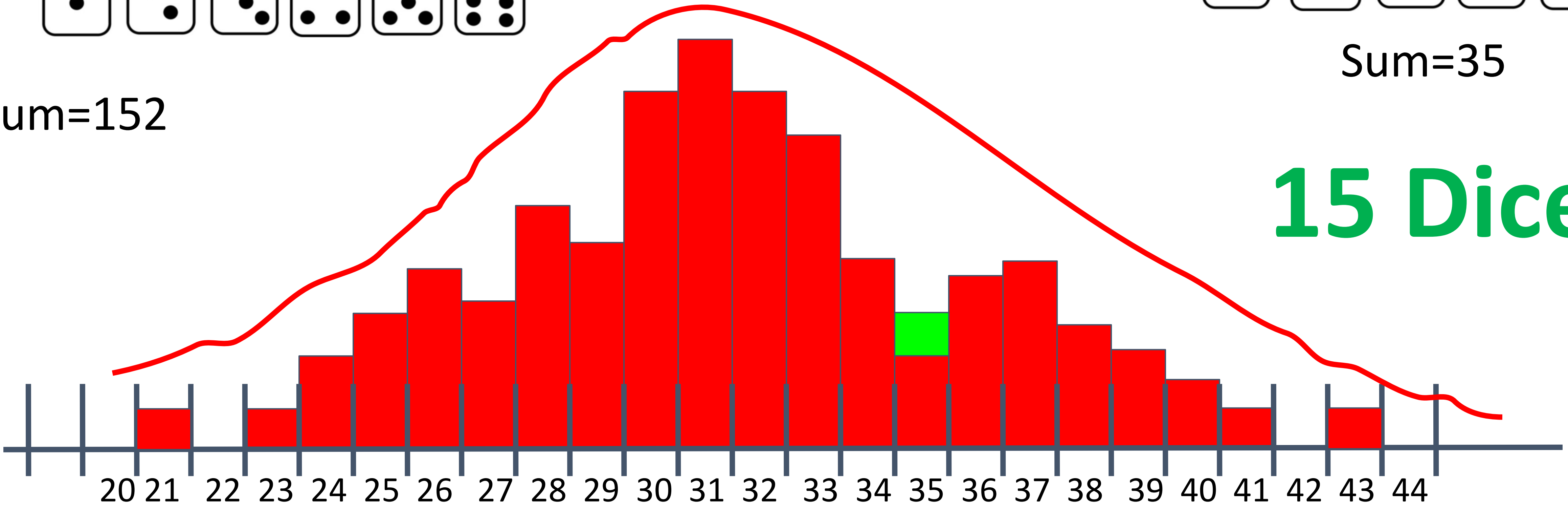
Example





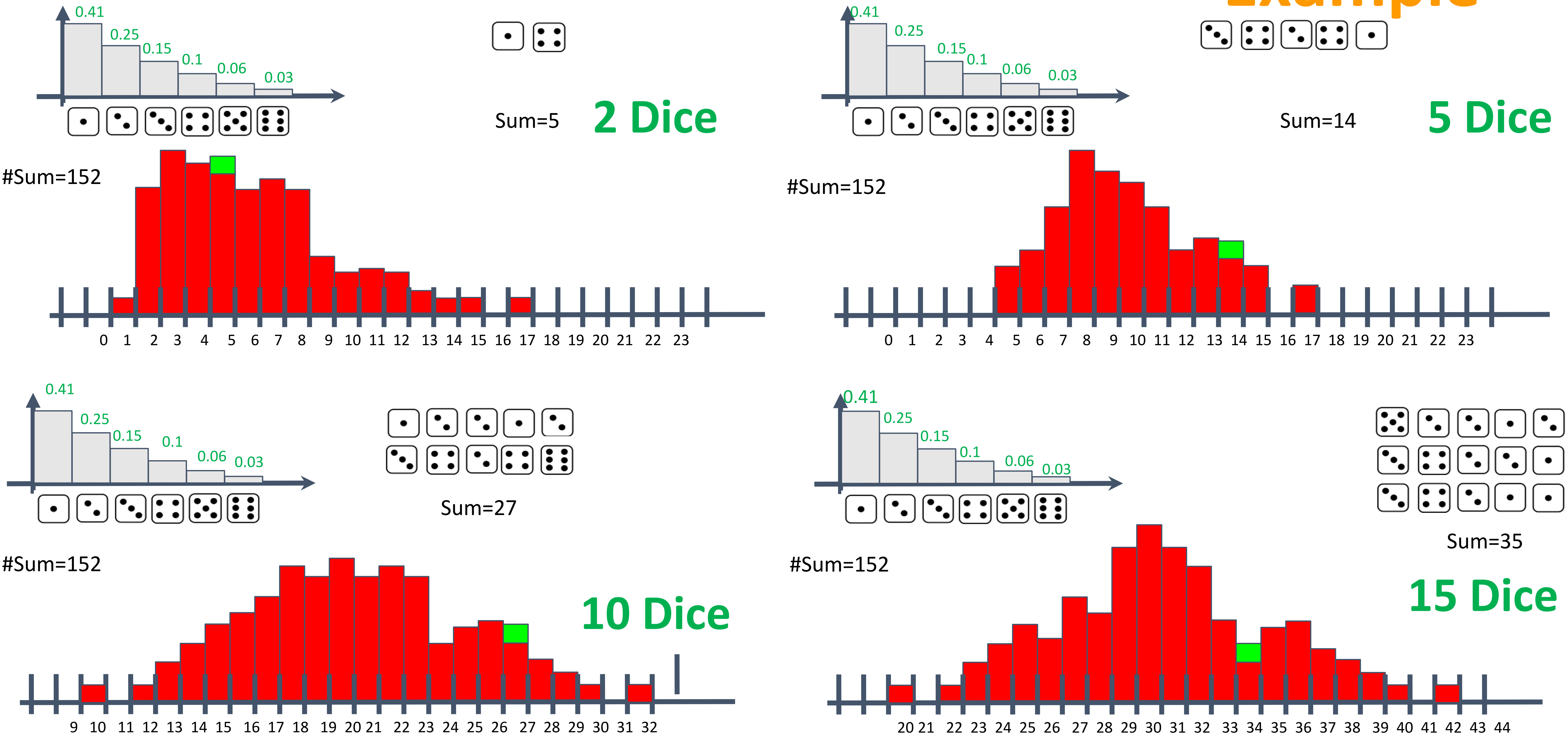
Sum=35

#Sum=152



15 Dice

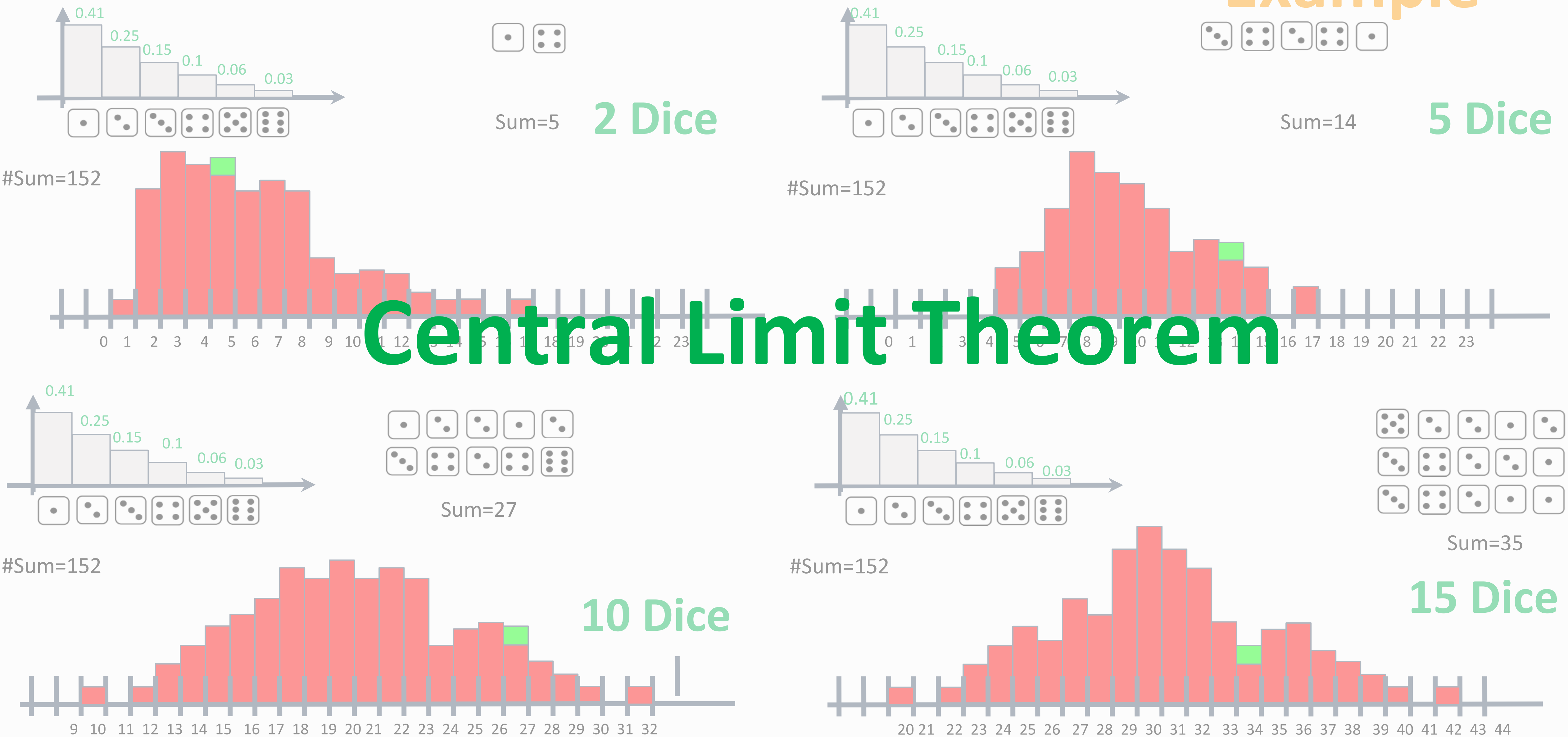
Example



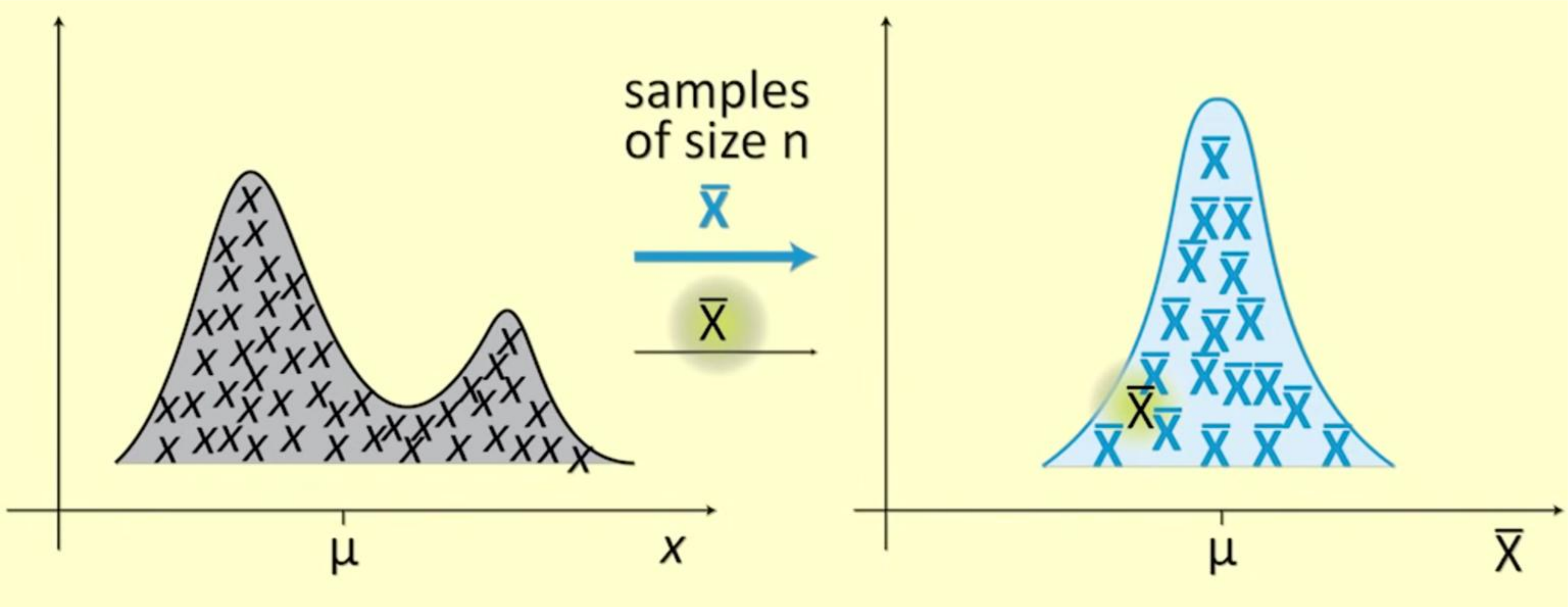
0.41
0.25
0.15
0.1
0.06
0.03
Sum=5
2 Dice
#Sum=152
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23
0.41
0.25
0.15
0.1
0.06
0.03
Sum=14
5 Dice
#Sum=152
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23
0.41
0.25
0.15
0.1
0.06
0.03
Sum=27
#Sum=152
10 Dice
9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32
0.41
0.25
0.15
0.1
0.06
0.03
Sum=35
#Sum=152
15 Dice
20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44

Example

Central Limit Theorem



Definition



Original Distribution

Sampling Distribution of the mean

Definition

Suppose $X_1, X_2, \ldots, X_n, \ldots$ are i.i.d. random variables each having mean μ and standard deviation σ. For each n let S_n denote the sum and let $\overline{X}_n$ be the average of $X_1, \ldots, X_n$.

$$\boxed{S_n = X_1 + X_2 + \ldots + X_n = \sum_{i=1}^{n} X_i}$$

$$\boxed{\overline{X}_n = \frac{X_1 + X_2 + \ldots + X_n}{n} = \frac{S_n}{n}.}$$

The properties of mean and variance show

$$\boxed{E(S_n) = n\mu, \qquad \mathrm{Var}(S_n) = n\sigma^2, \qquad \sigma_{S_n} = \sqrt{n}\,\sigma}$$

$$\boxed{E(\overline{X}_n) = \mu, \qquad \mathrm{Var}(\overline{X}_n) = \frac{\sigma^2}{n}, \qquad \sigma_{\overline{X}_n} = \frac{\sigma}{\sqrt{n}}.}$$

Since they are multiples of each other, S_n and $\overline{X}_n$ have the same standardization

$$Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}} = \frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}}$$

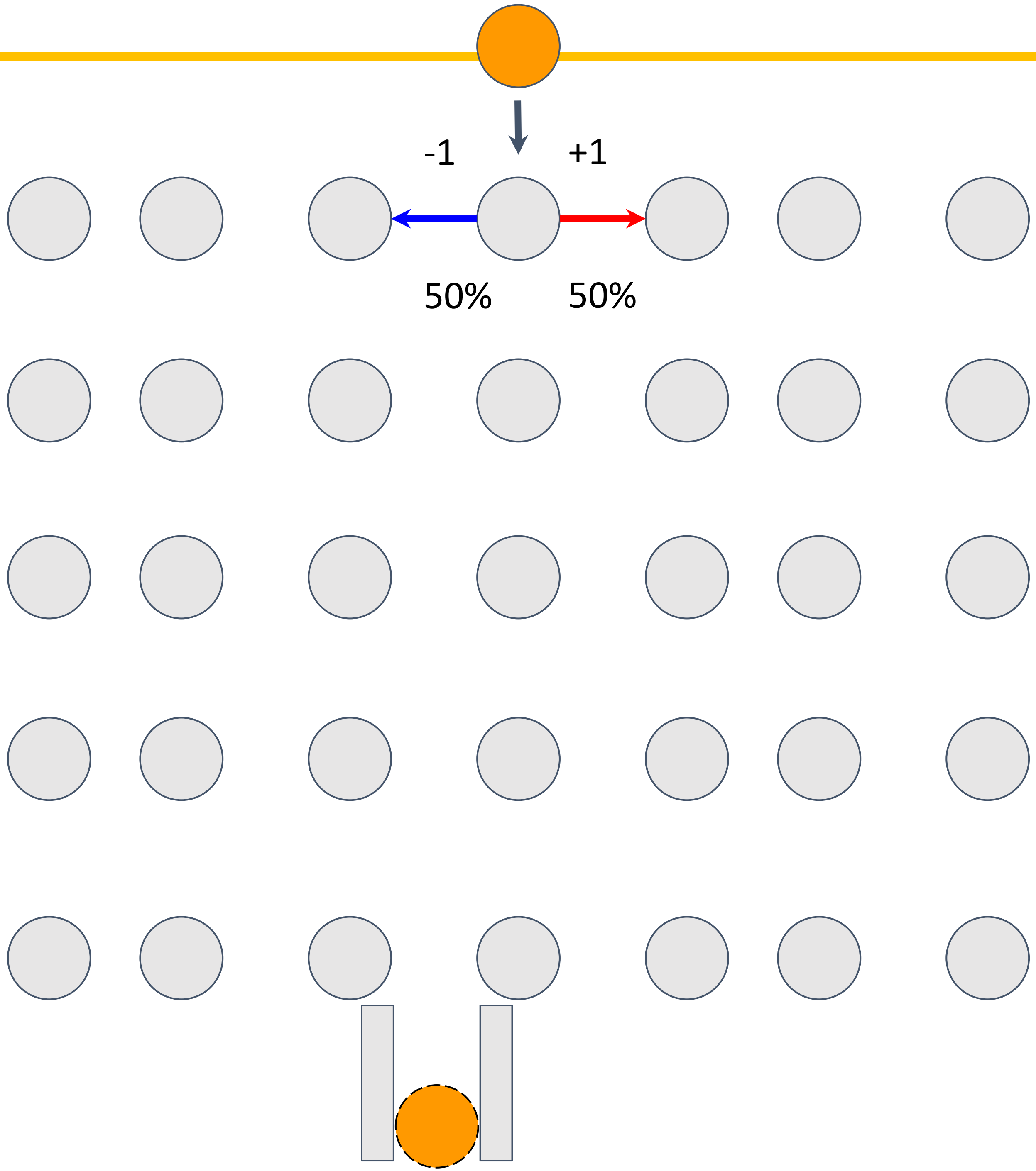
Central Limit Theorem: For large n,

$$\overline{X}_n \approx \mathrm{N}(\mu, \sigma^2/n), \qquad S_n \approx \mathrm{N}(n\mu, n\sigma^2), \qquad Z_n \approx \mathrm{N}(0,1).$$

Example

Galton Board

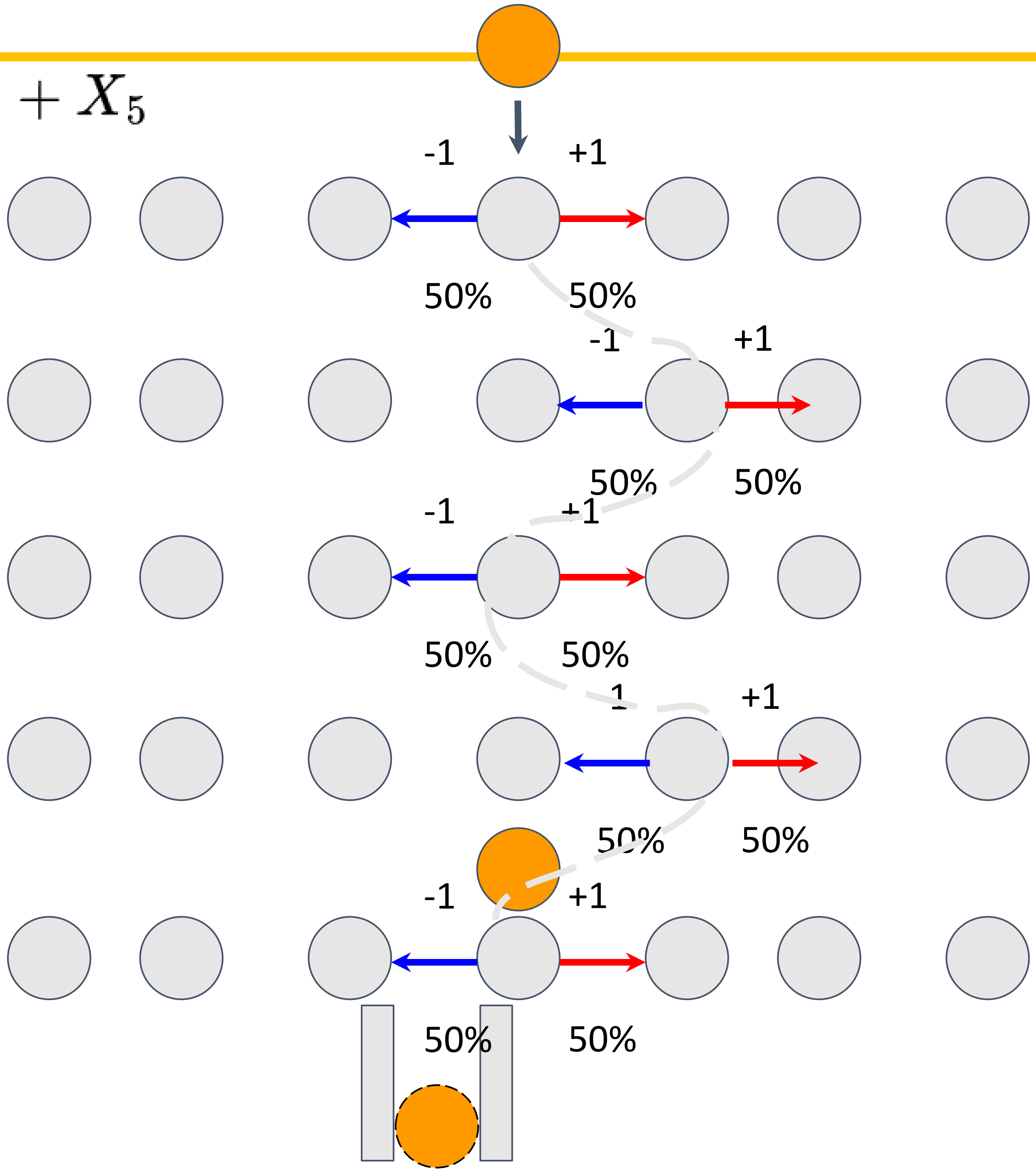
Explanation



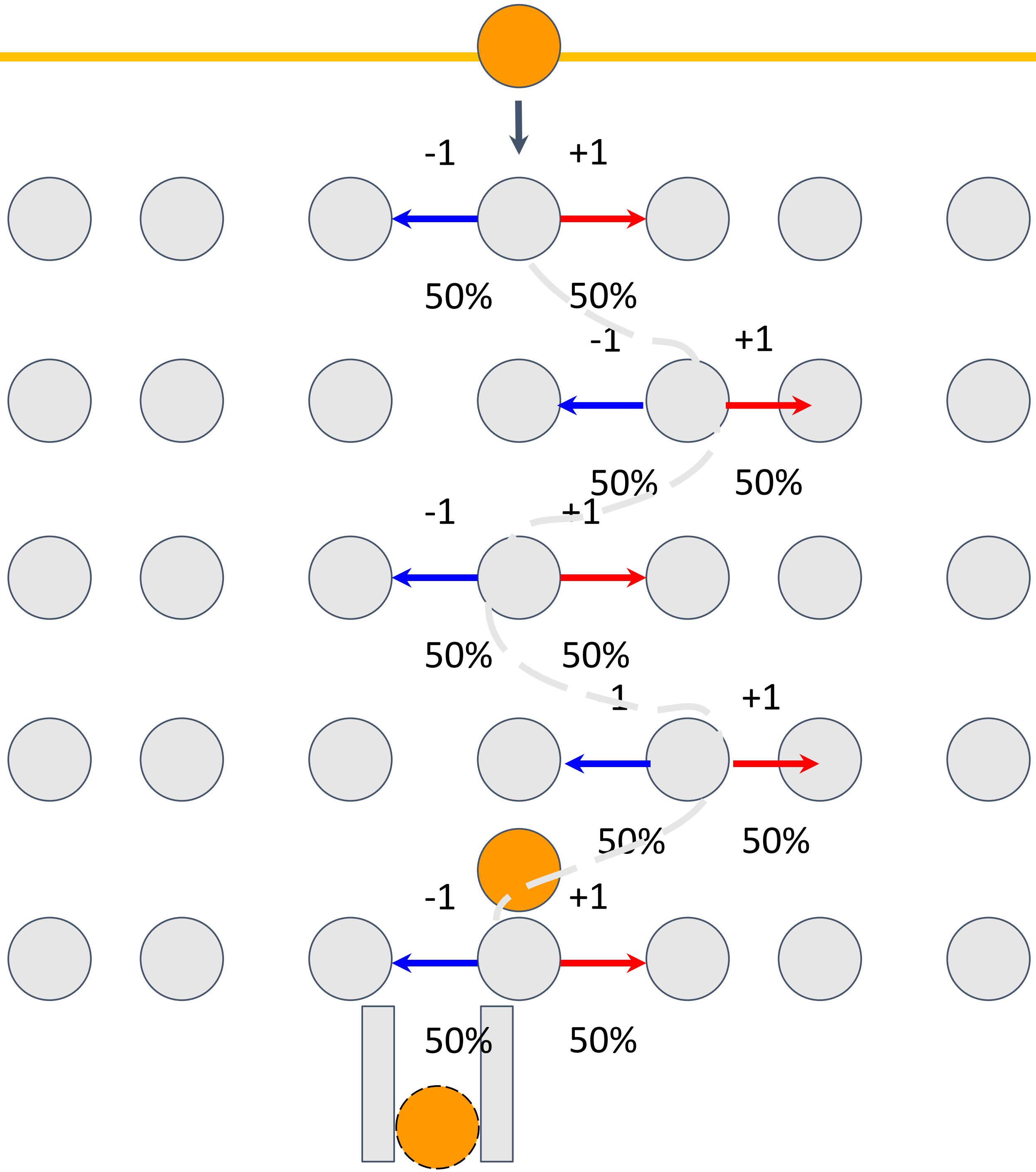
Explanation

$$X_1 + X_2 + X_3 + X_4 + X_5$$

+1 - 1 - 1 + 1 -1 = -1



Explanation



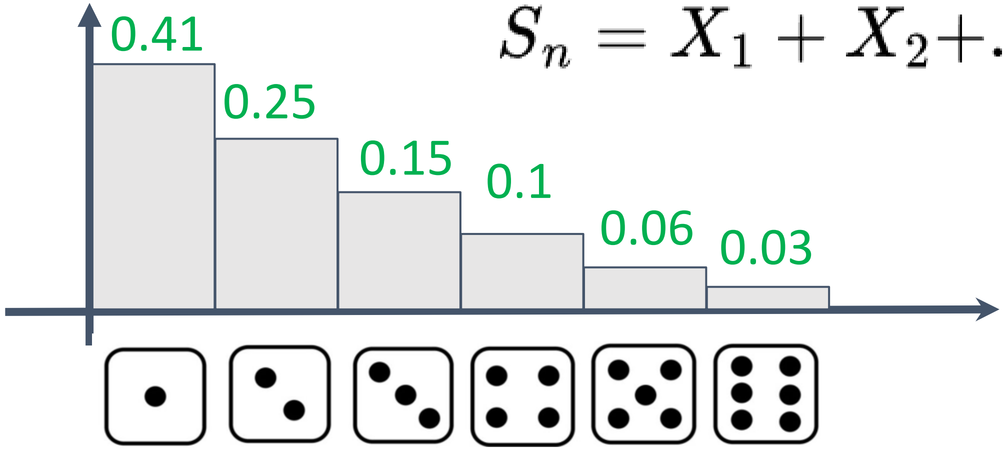
$$X_1 + X_2 + X_3 + X_4 + X_5$$

+1 - 1 - 1 + 1 -1 = -1

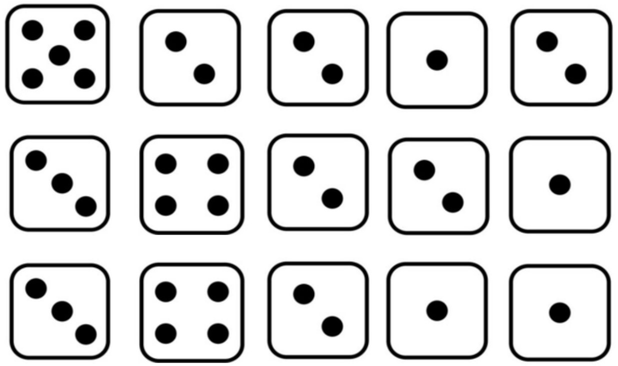
$$S_n = X_1 + X_2 + X_3 + X_4 + X_5 \approx \mathcal{N}(5\mu, 5\sigma^2)$$

Explanation

$$S_n = X_1 + X_2 + \ldots + X_{15} \approx \mathcal{N}(15\mu, 15\sigma^2)$$



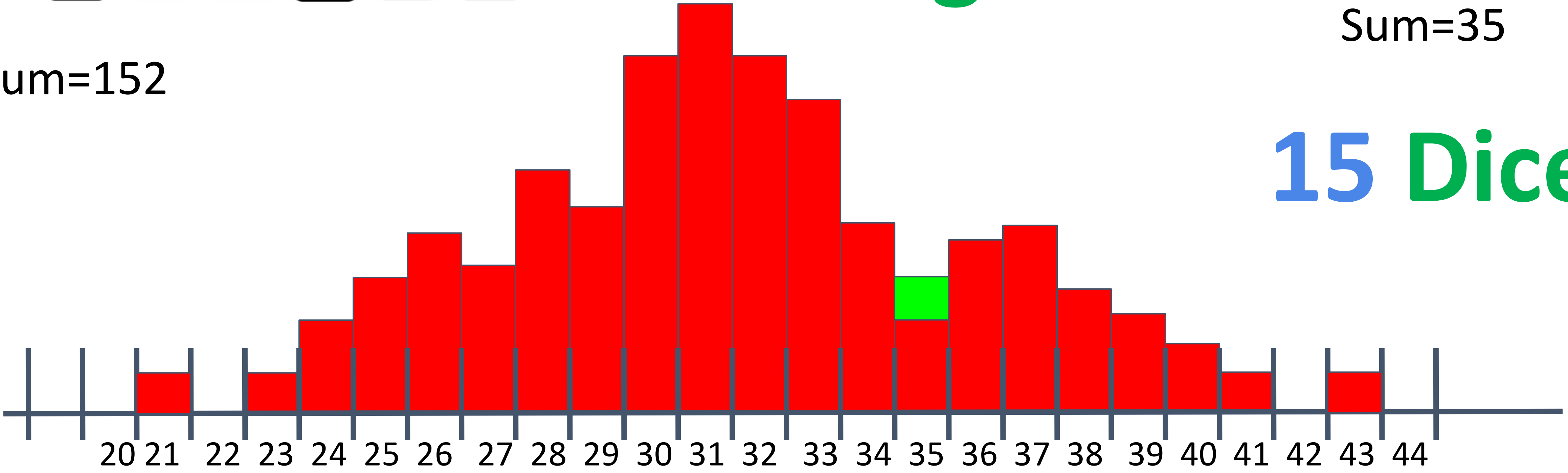
n large enough



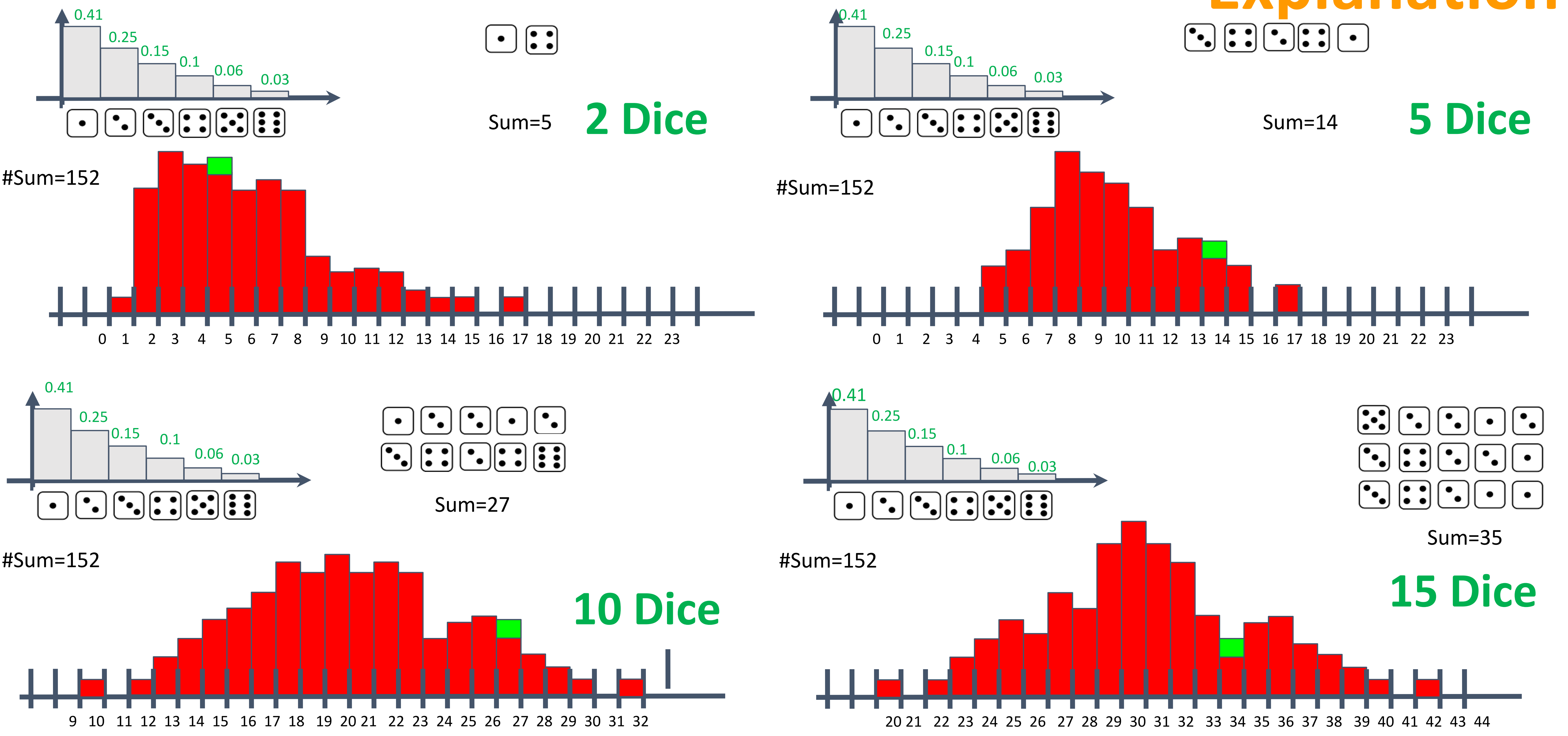
Sum=35

#Sum=152

15 Dice



Explanation



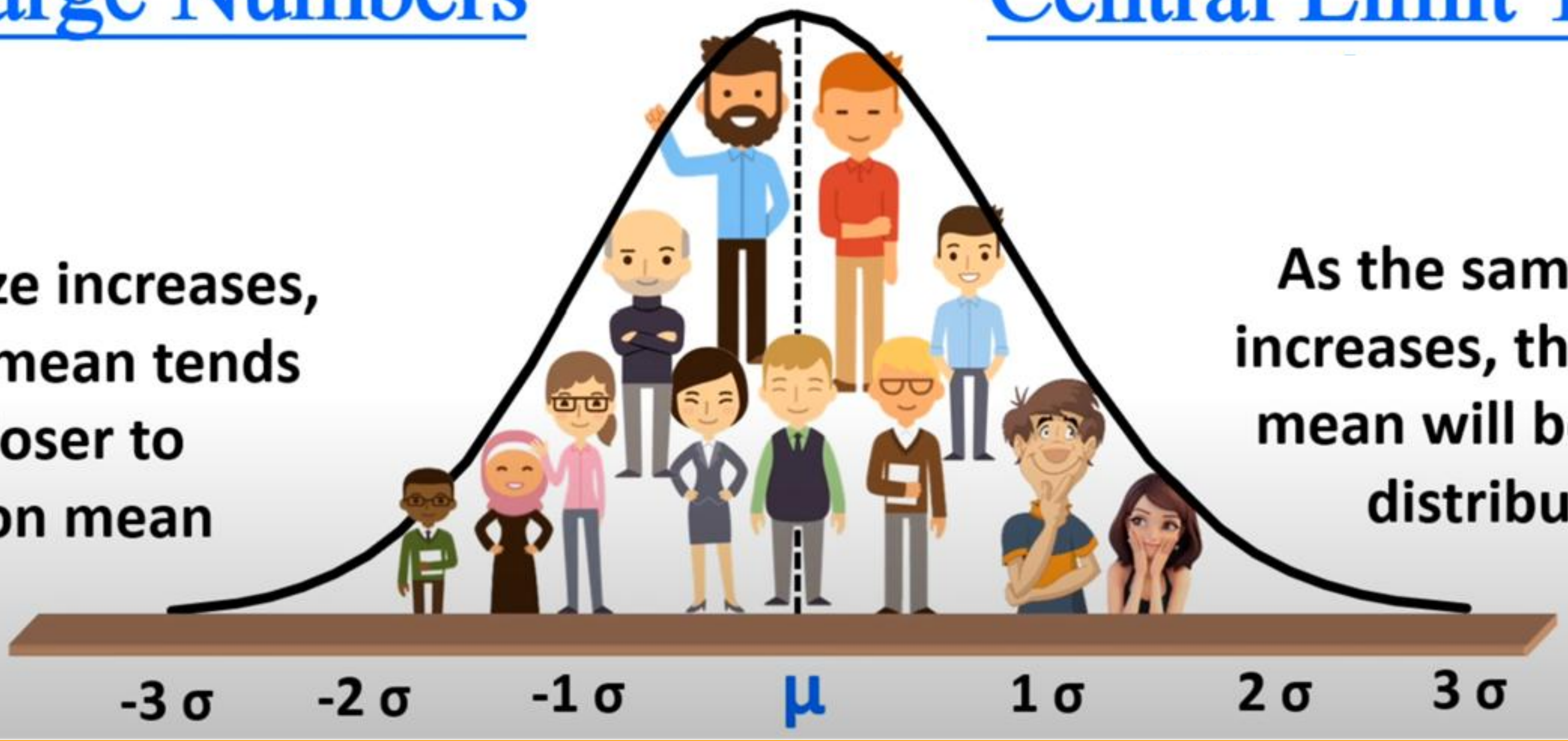
Comparison

Law of Large Numbers

As Sample size increases, the Sample mean tends to get closer to Population mean

Central Limit Theorem

As the sample size increases, the sample mean will be evenly distributed.



Conditions for the Law of Large Numbers And Central Limit Theorem

- Independent identically distributed variables
- Finite mean and finite variances
- Large sample size

Questions?