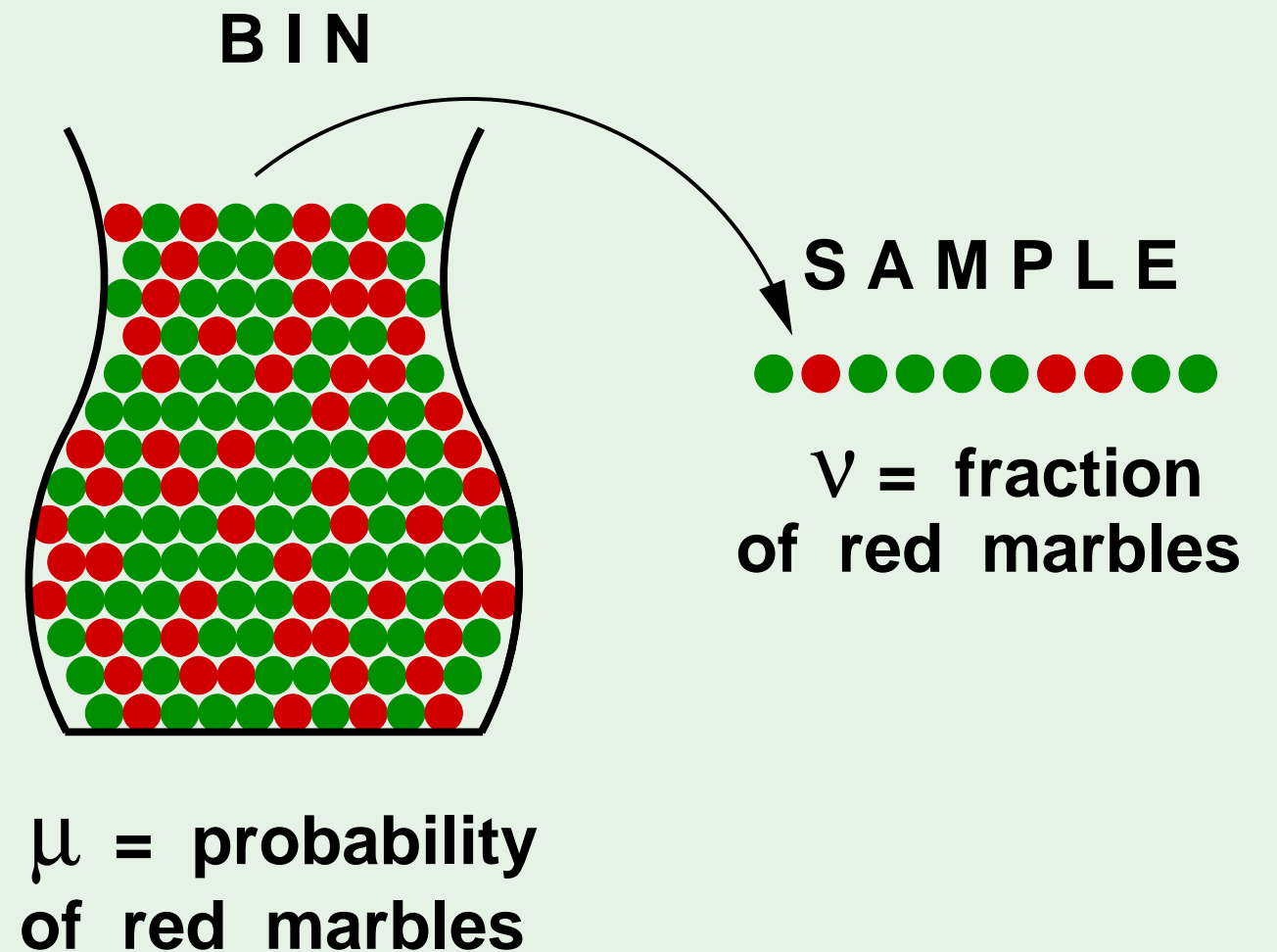# A related experiment

- Consider a 'bin' with red and green marbles.

$$\mathbb{P}[\text{ picking a red marble }] = \mu$$

$$\mathbb{P}[\text{ picking a green marble }] = 1 - \mu$$

- The value of μ is <u>unknown</u> to us.

- We pick N marbles independently.

- The fraction of red marbles in sample $= \nu$

B I N

S A M P L E

ν = fraction of red marbles

μ = probability of red marbles
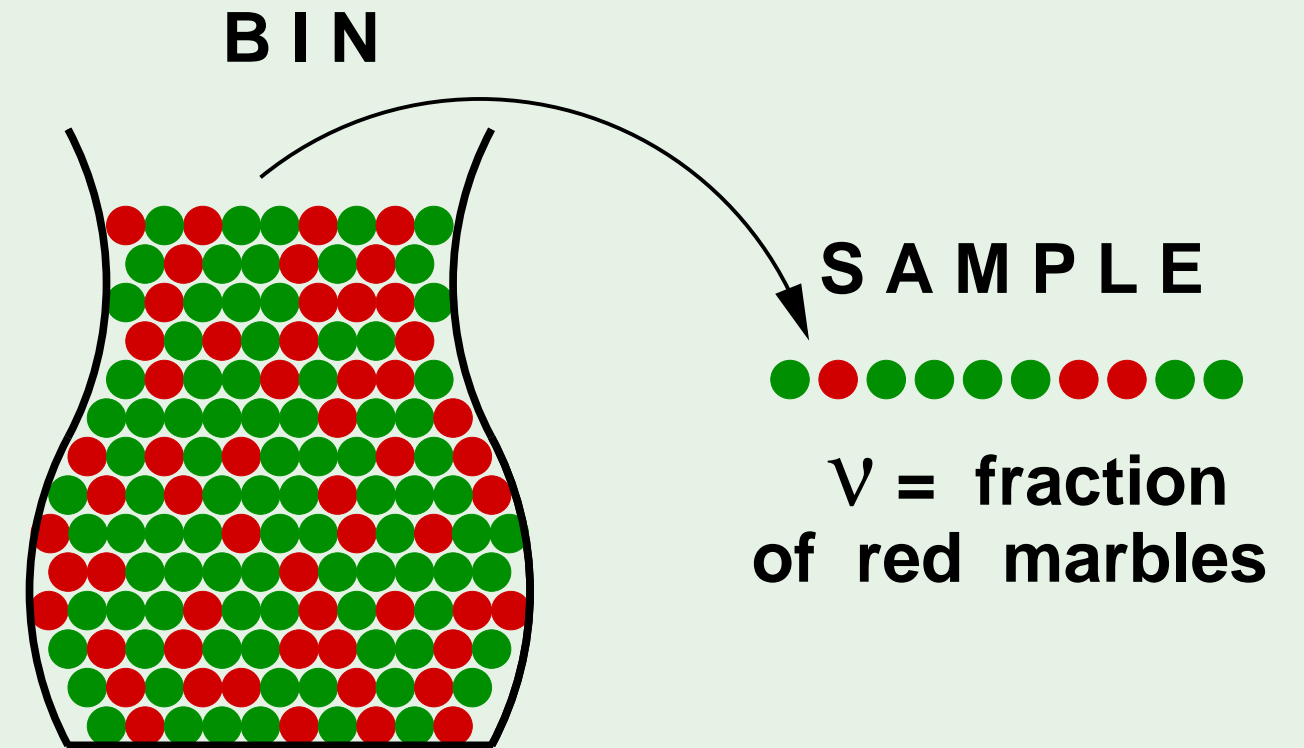
Does ν say anything about μ?

No!

Sample can be mostly green while bin is mostly red.

Yes!

Sample frequency ν is likely close to bin frequency μ.

possible versus **probable**

B I N

S A M P L E

ν = **fraction of red marbles**

μ = **probability of red marbles**

What <u>does</u> ν say about μ?

In a big sample (large N), ν is probably close to μ (within ϵ).

Formally,

$$\mathbb{P}\left[\left|\nu - \mu\right| > \epsilon\right] \leq 2e^{-2\epsilon^2 N}$$

This is called **Hoeffding's Inequality**.

In other words, the statement "$\mu = \nu$" is P.A.C.