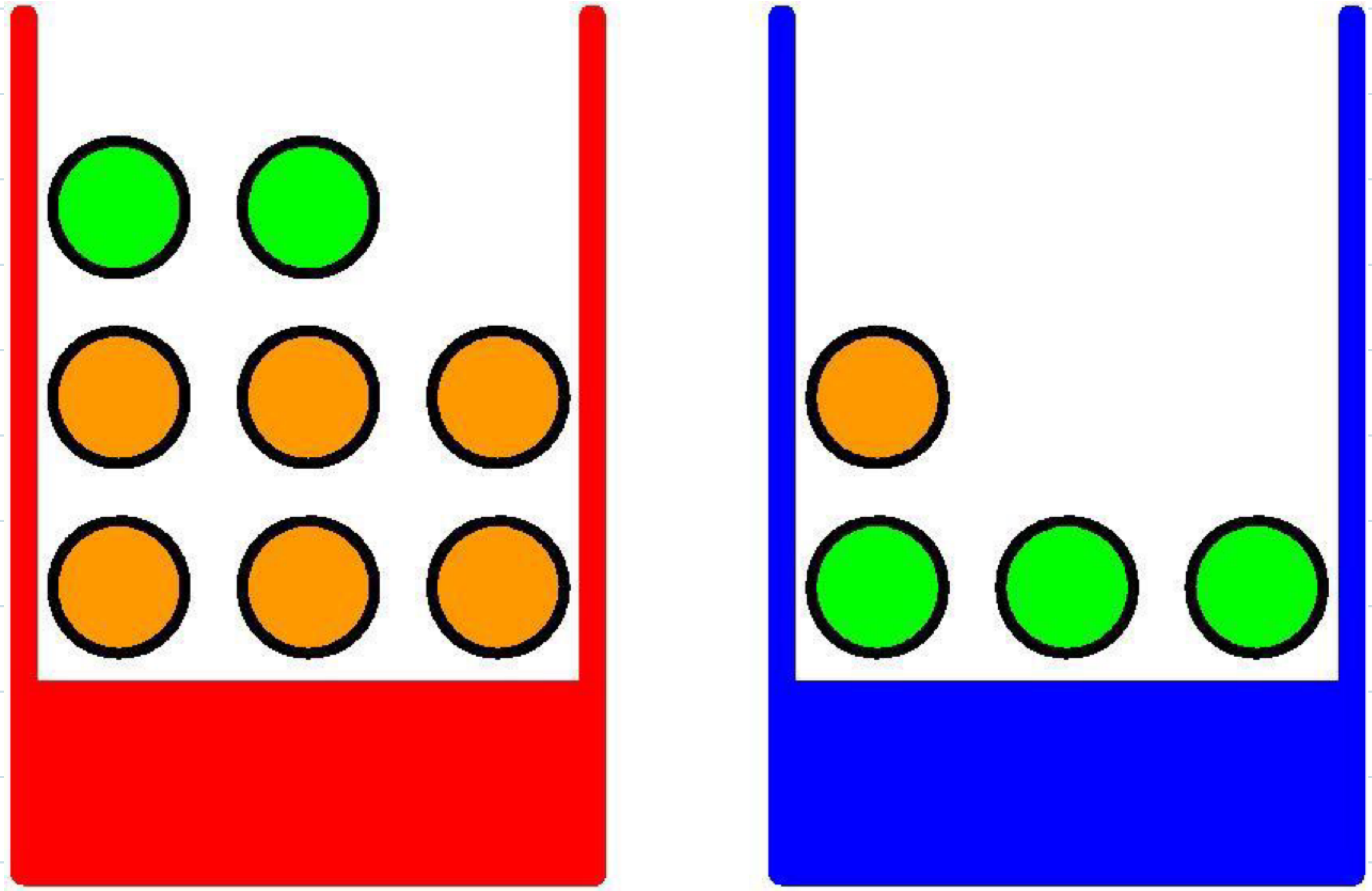
Probability Theory (PT)

Notes:

1) PT gives us a framework for quantifying uncertainty in our models and their predictions.
2) The probability of an event is the fraction of time an event occurs as the number of trials goes to infinity.

Example: Apples and oranges

Two random variables



B = red or blue

F = apple or orange

Let's say

$$\boxed{\begin{aligned} p(B=\text{red}) &= 0.4 \\ p(B=\text{blue}) &= 0.6 \end{aligned}}$$

The probabilities add to 1.

N is the total counts in all the blocks

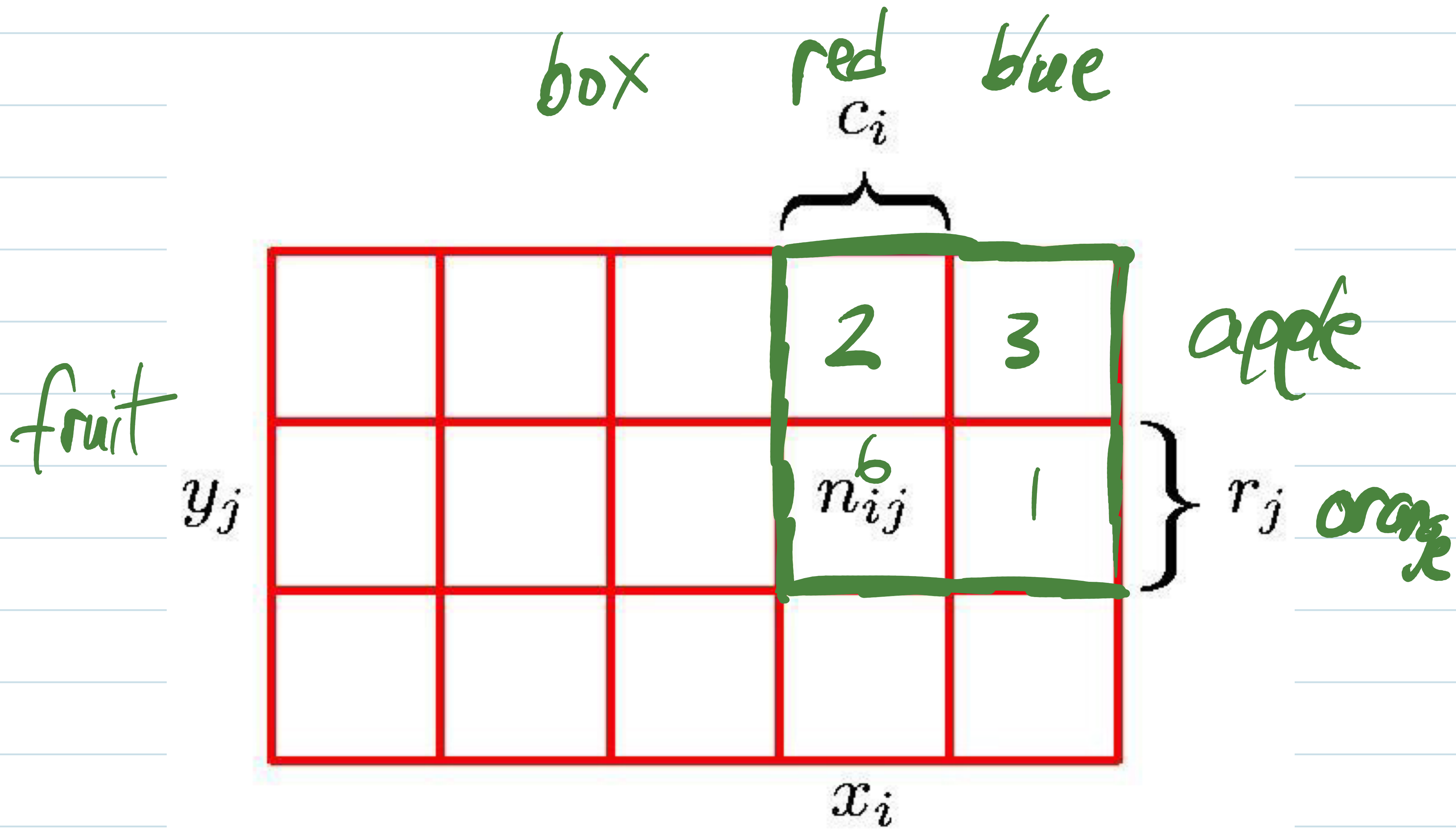
n_{ij} is the counts in block ij

c_i Counts in all blocks of column i

r_j Counts in all blocks of row j

X random variable that takes on a value x_i

Y random variable that takes on a value y_j



Marginal probability

$$p(X = x_i) = \frac{c_i}{N}$$

e.g $p(B = red) = \frac{8}{12}$

Joint probability

$$p(X = x_i, Y = y_i) = \frac{n_{ij}}{N}$$

e.g $p(B = red, F = apple) = \frac{2}{12}$

Conditional Probability

$$p(Y = y_j \mid X = x_i) = \frac{n_{ij}}{c_i}$$

e.g $p(F = apple \mid B = red) = \frac{2}{8}$

sum rule $\quad p(X=x_i) = \dfrac{c_i}{N} = \dfrac{1}{N}\sum_{j=1}^{L} n_{ij}$

product rule $\quad p(X=x_i, Y=y_j) = \dfrac{n_{ij}}{c_i} * \dfrac{c_i}{N}$

$$= p(Y=y_j \mid X=x_i)\,p(X=x_i)$$

e.g $\quad p(B=\text{red}, F=\text{orange}) = p(F=o \mid B=\text{red})\,p(B=\text{red})$

$$0.75 * 0.40 = 0.30$$

Sum Rule $$p(X) = \sum_{Y} p(X,Y)$$

Product Rule $$p(X,Y) = p(Y|X)\,p(X)$$

Bayes' Theorem $$p(Y|X) = \frac{p(X|Y)\,p(Y)}{p(X)}$$

$$p(X) = \sum_{Y} p(X|Y)\,p(Y)$$

posterior likelihood prior