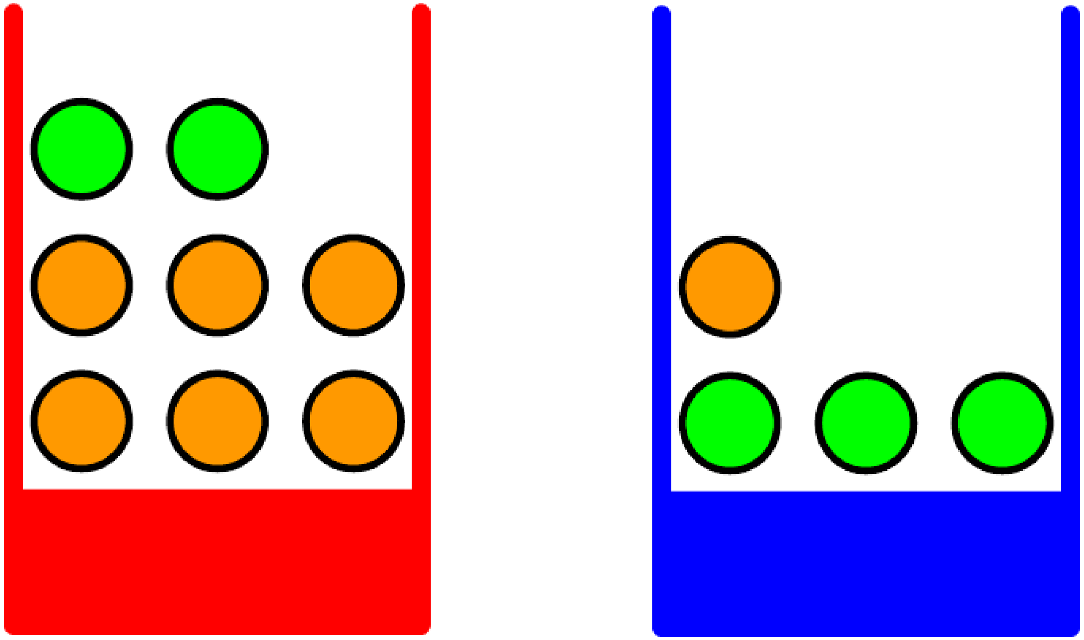
Apples and oranges

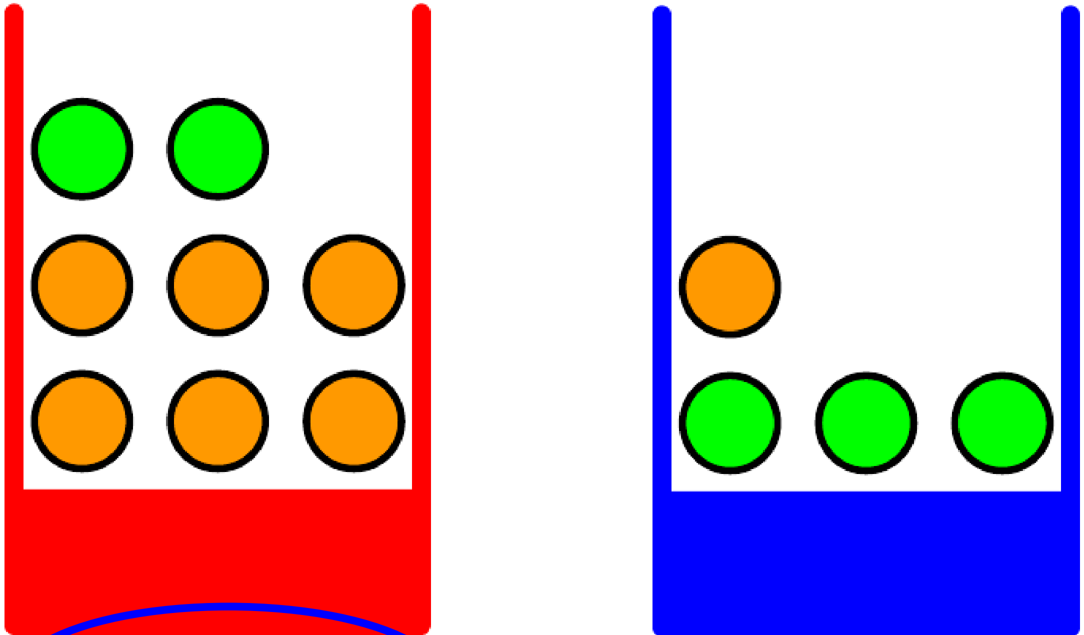
– Randomly choose one box, and then randomly pick a piece of fruit from that box



$p(B = r) = 4/10$ $\quad$ $p(B = b) = 6/10$

Apples and oranges

– Randomly choose one box, and then randomly pick a piece of fruit from that box



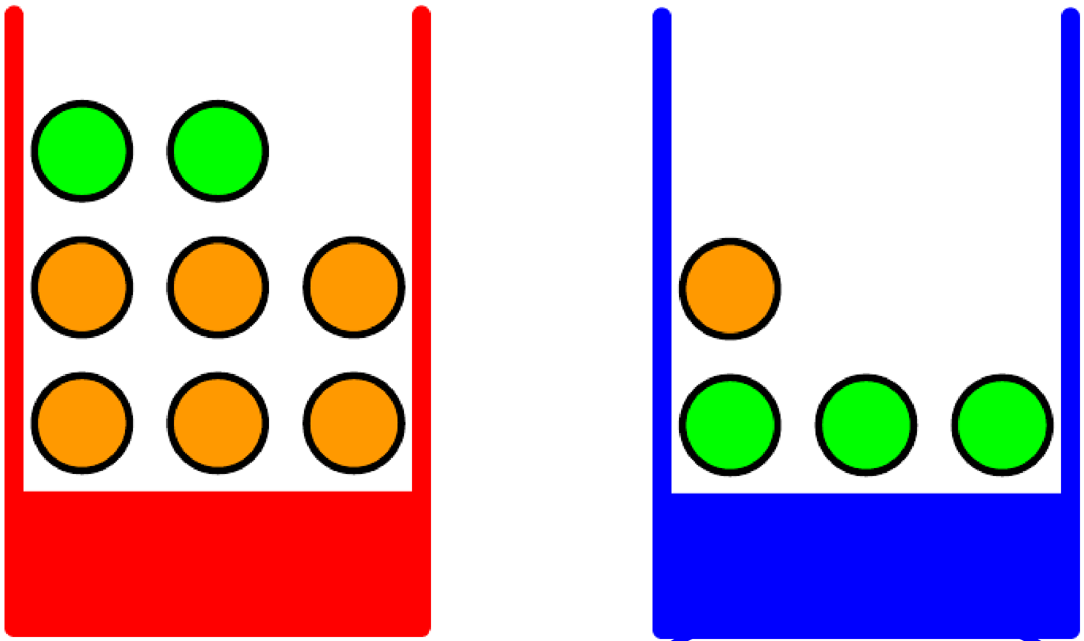
$p(B = r) = 4/10$ $\quad$ $p(B = b) = 6/10$

probability we

choose red Box

Apples and oranges

- Randomly choose one box, and then randomly pick a piece of fruit from that box



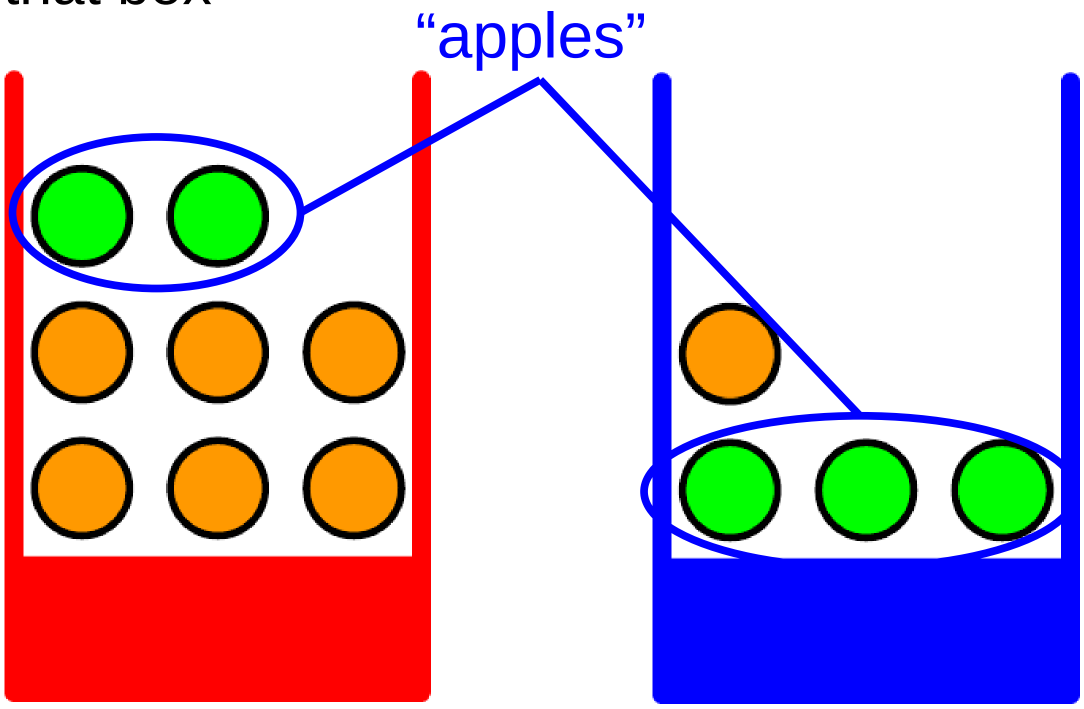
$p(B = r) = 4/10$

$p(B = b) = 6/10$

probability we choose <u>b</u>lue <u>B</u>ox

Apples and oranges

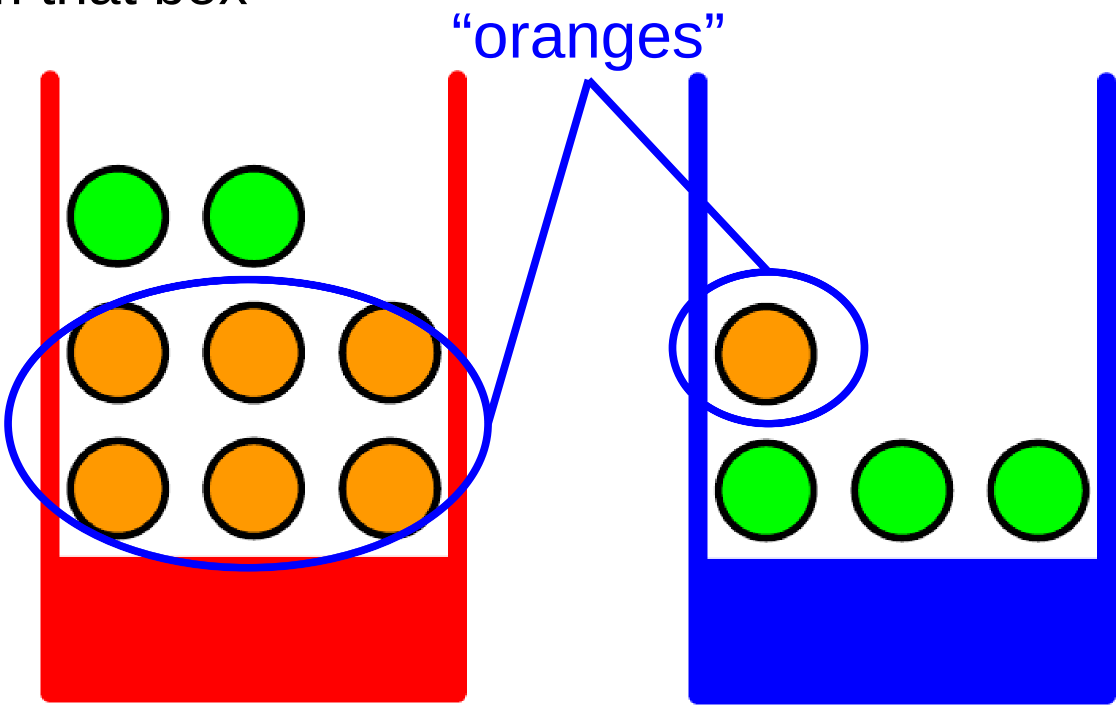
– Randomly choose one box, and then randomly pick a piece of fruit from that box



$p(B = r) = 4/10$ $p(B = b) = 6/10$

Apples and oranges

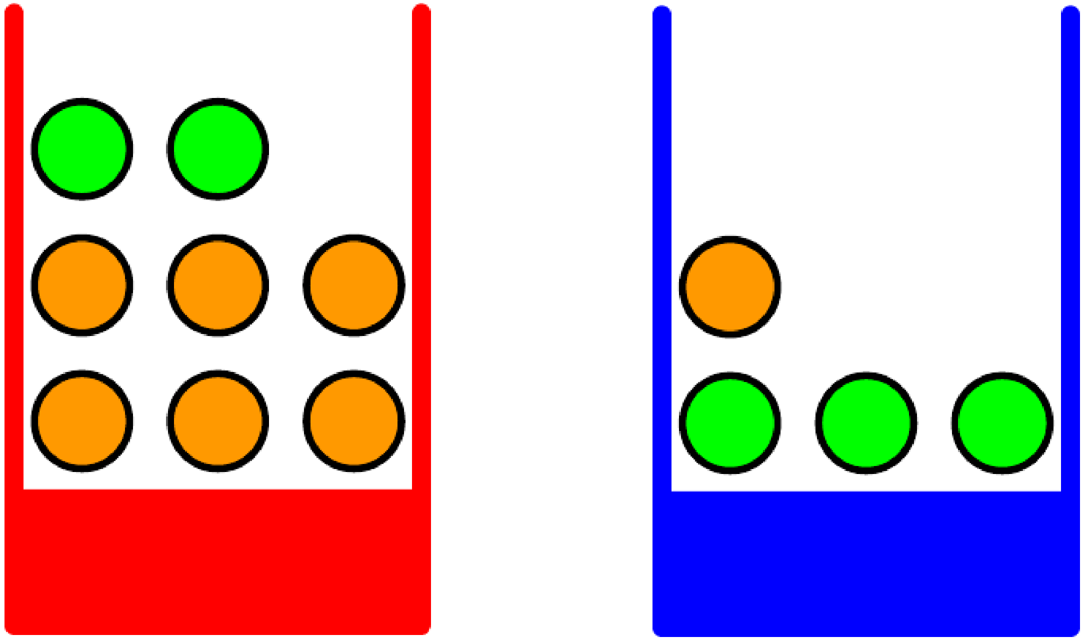
– Randomly choose one box, and then randomly pick a piece of fruit from that box



$p(B = r) = 4/10$ $\quad$ $p(B = b) = 6/10$

Apples and oranges

– Randomly choose one box, and then randomly pick a piece of fruit from that box



$p(B = r) = 4/10$ $\quad$ $p(B = b) = 6/10$

(1) What is the probability we get an apple?

(2) If the fruit we end up with is an orange, what is the probability we had chosen the red box?

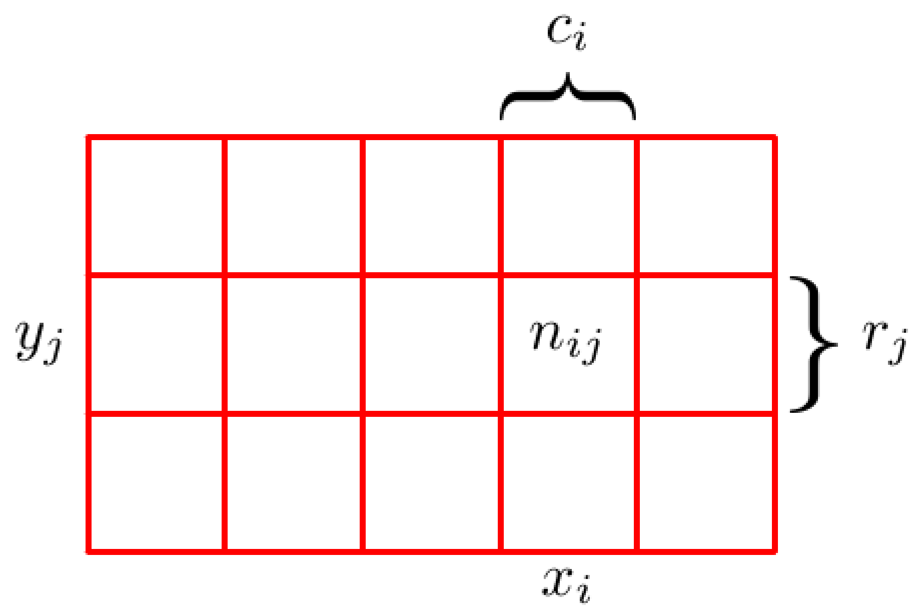
The two rules of probability

- Two random variables:

  X takes values $\{x_i\}$ where $i = 1, \ldots, M$

  Y takes values $\{y_j\}$ where $j = 1, \ldots, L$

- Observe outcomes (X, Y) of N samples, with $N \rightarrow \infty$



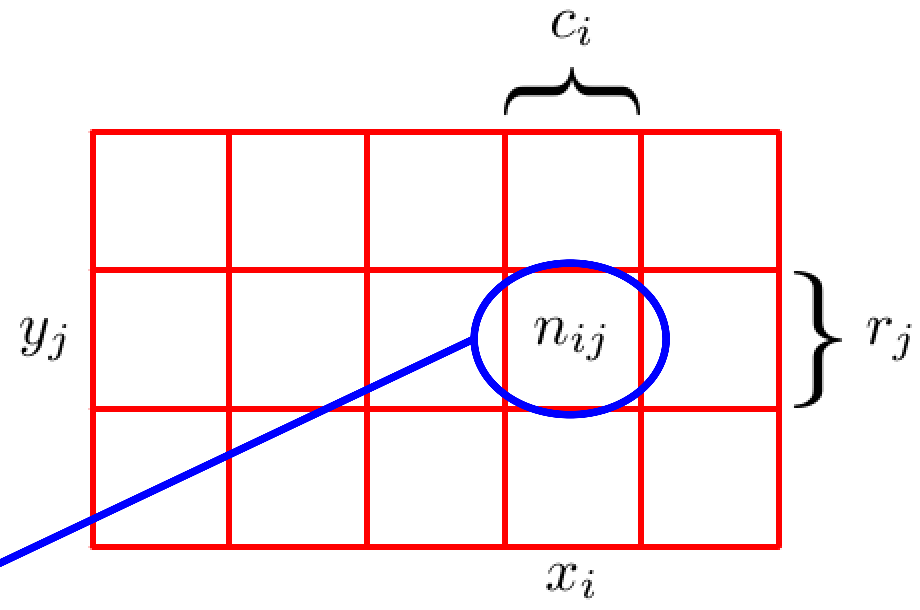
The two rules of probability

- Two random variables:

  X takes values $\{x_i\}$ where $i = 1, \ldots, M$

  Y takes values $\{y_j\}$ where $j = 1, \ldots, L$

- Observe outcomes (X, Y) of N samples, with $N \to \infty$



number of samples in which $X = x_i$ and $Y = y_j$

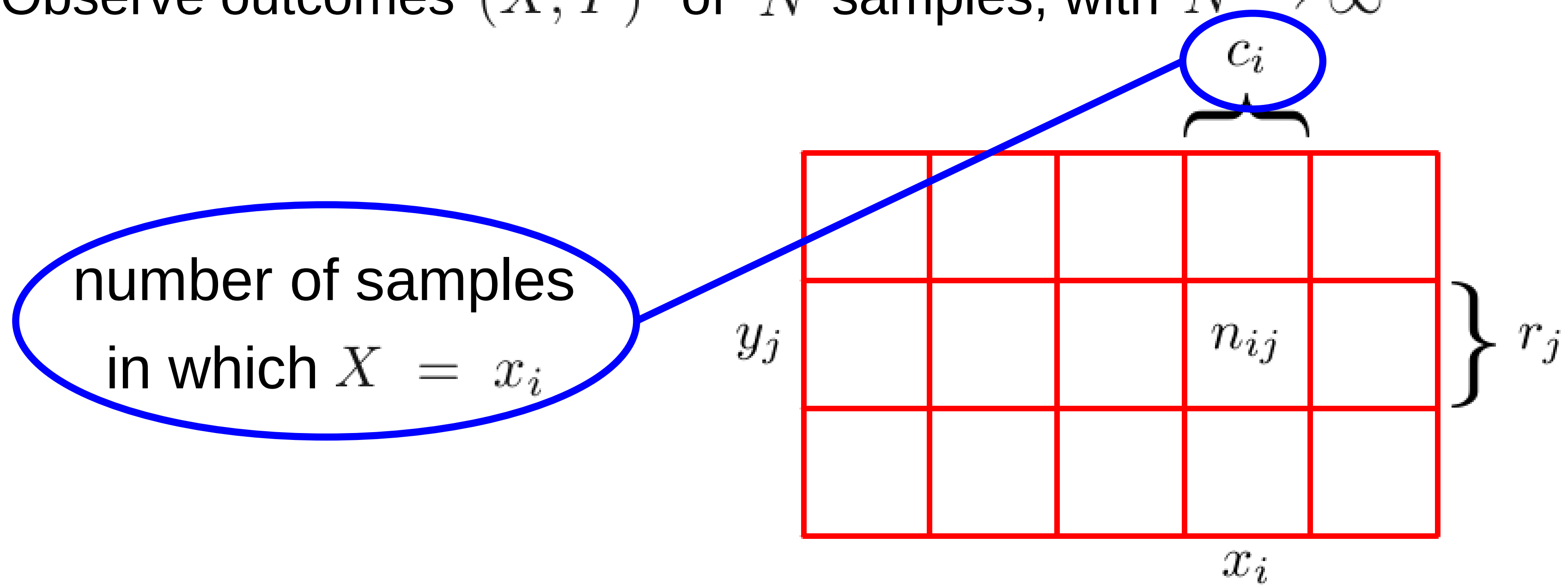
The two rules of probability

– Two random variables:

X takes values $\{x_i\}$ where $i = 1, \ldots, M$

Y takes values $\{y_j\}$ where $j = 1, \ldots, L$

– Observe outcomes (X, Y) of N samples, with $N \rightarrow \infty$



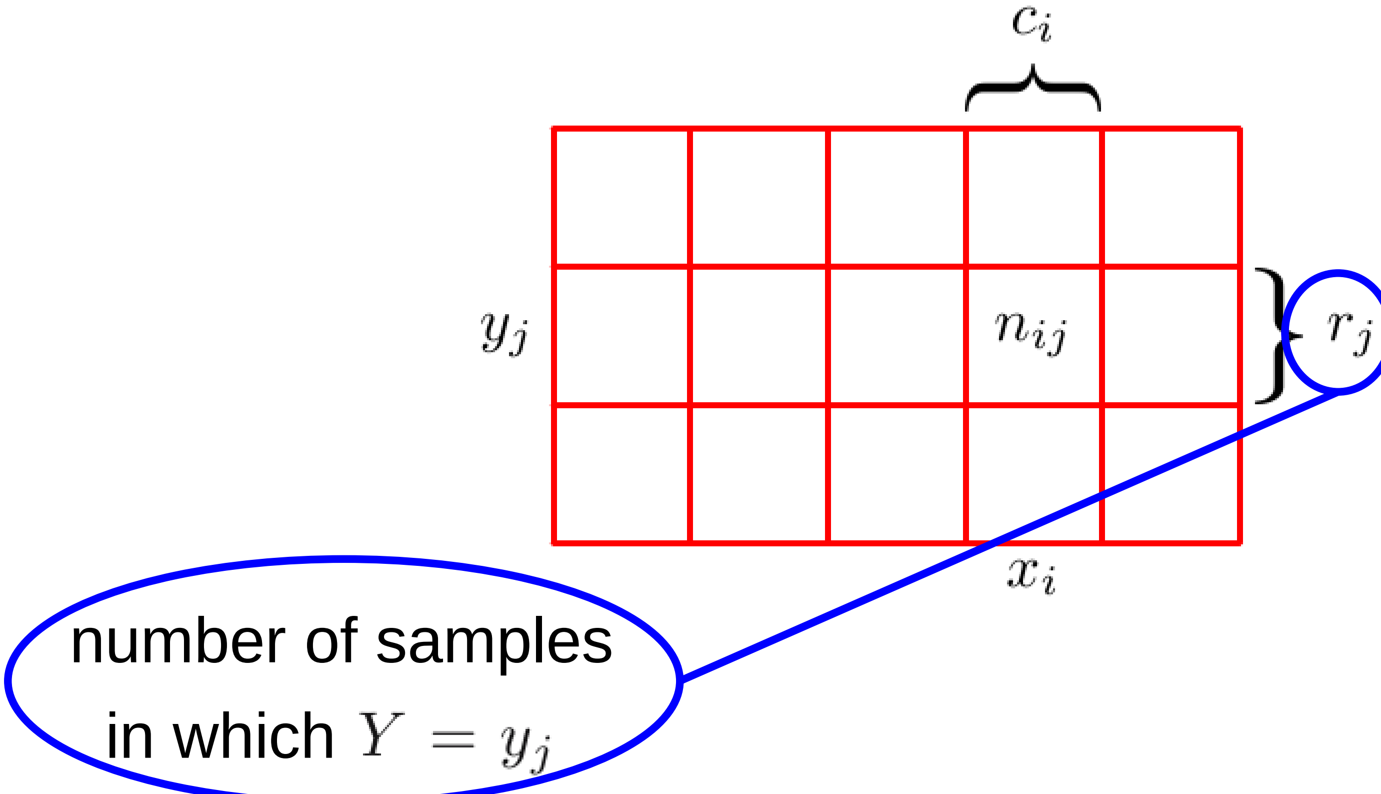
The two rules of probability

- Two random variables:

  X takes values $\{x_i\}$ where $i = 1, \ldots, M$

  Y takes values $\{y_j\}$ where $j = 1, \ldots, L$

- Observe outcomes (X, Y) of N samples, with $N \to \infty$



The two rules of probability

- Two random variables:

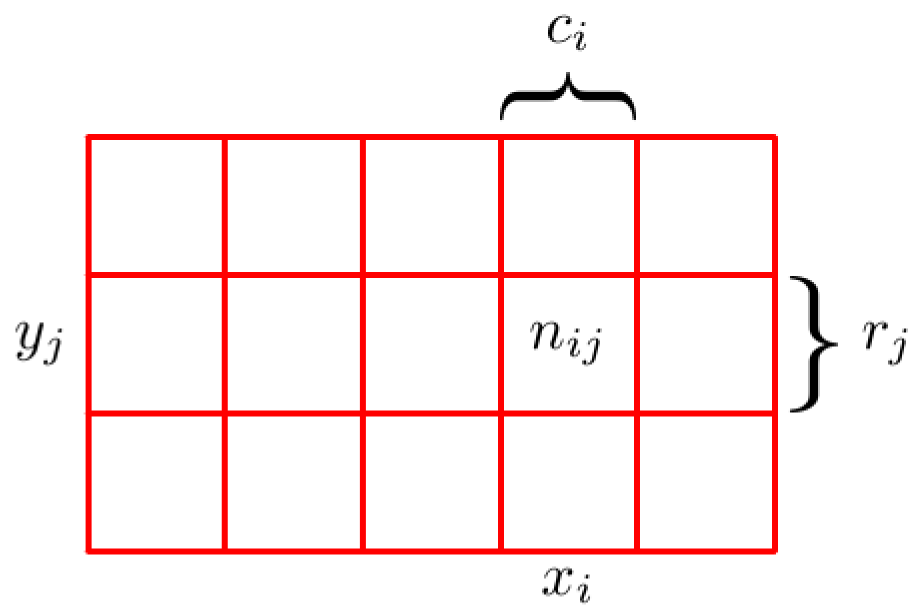
  X takes values $\{x_i\}$ where $i = 1, \ldots, M$

  Y takes values $\{y_j\}$ where $j = 1, \ldots, L$

- Observe outcomes (X, Y) of N samples, with $N \rightarrow \infty$

"joint probability"

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



The two rules of probability

- Two random variables:

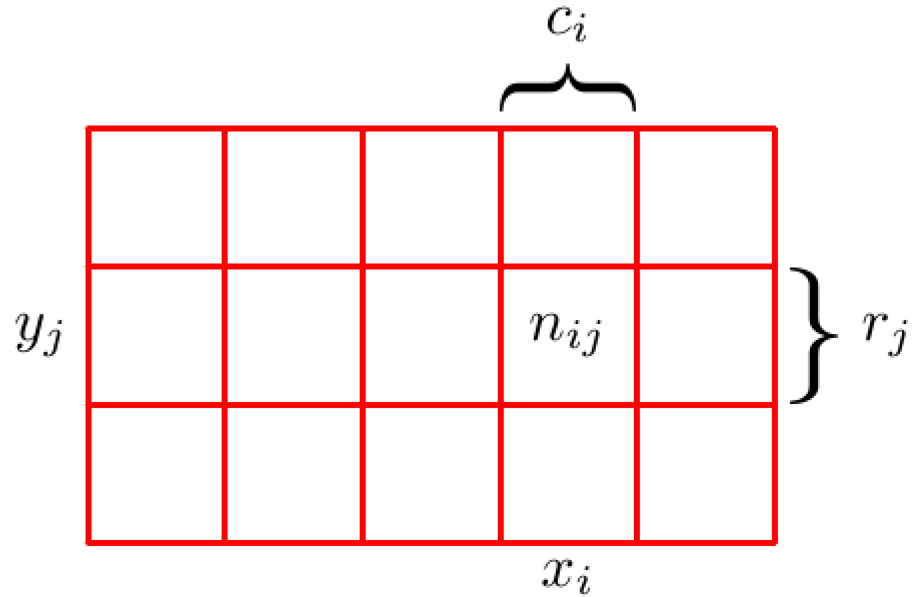
  X takes values $\{x_i\}$ where $i = 1, \ldots, M$

  Y takes values $\{y_j\}$ where $j = 1, \ldots, L$

- Observe outcomes (X, Y) of N samples, with $N \to \infty$

“marginal probability”

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



The two rules of probability

- Two random variables:

  X takes values $\{x_i\}$ where $i = 1, \ldots, M$

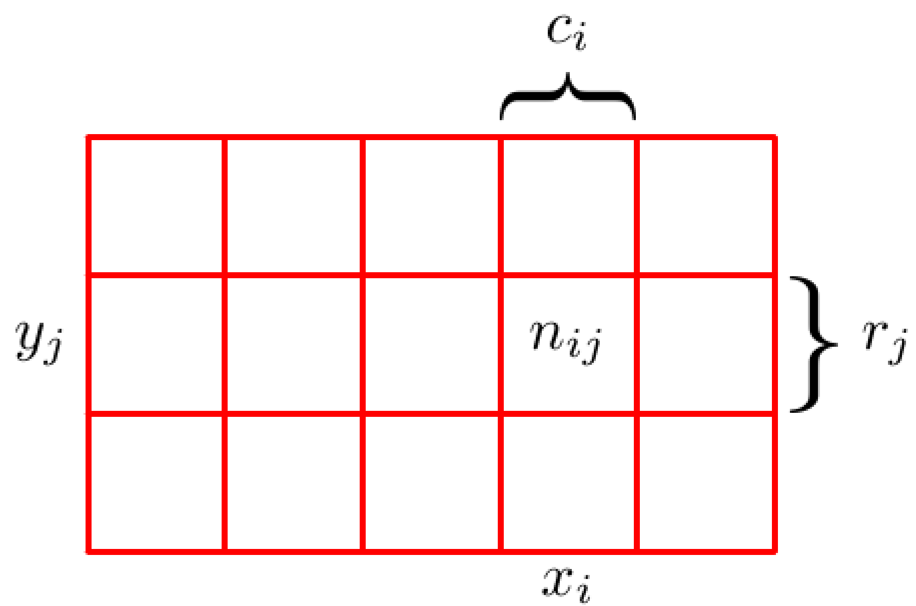
  Y takes values $\{y_j\}$ where $j = 1, \ldots, L$
- Observe outcomes (X, Y) of N samples, with $N \rightarrow \infty$

“marginal probability”

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

$$= \frac{\sum_j n_{ij}}{N}$$

$$= \sum_{j=1}^{L} p(X = x_i, Y = y_j)$$



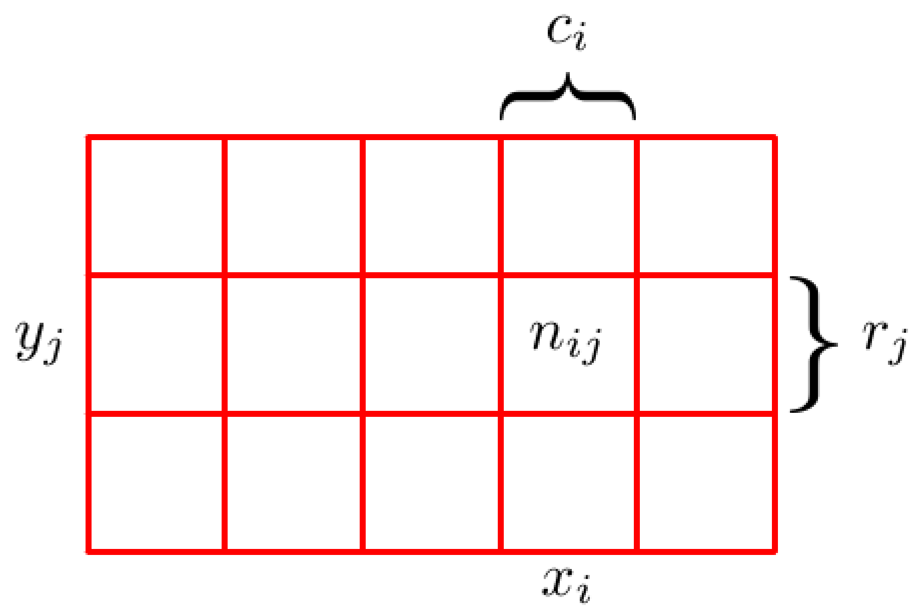
The two rules of probability

- Two random variables:

  X takes values $\{x_i\}$ where $i = 1, \ldots, M$

  Y takes values $\{y_j\}$ where $j = 1, \ldots, L$

- Observe outcomes (X, Y) of N samples, with $N \to \infty$



"marginal probability"

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

$$= \frac{\sum_j n_{ij}}{N}$$

$$= \sum_{j=1}^{L} p(X = x_i, Y = y_j)$$

"sum rule"

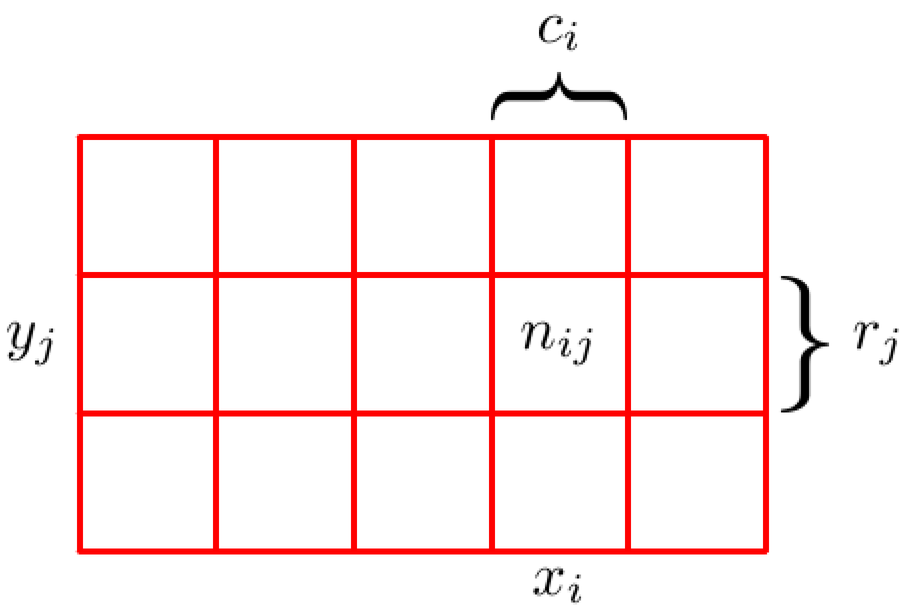
The two rules of probability

- Two random variables:

  X takes values $\{x_i\}$ where $i = 1, \ldots, M$

  Y takes values $\{y_j\}$ where $j = 1, \ldots, L$

- Observe outcomes (X, Y) of N samples, with $N \to \infty$



“conditional probability”

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

The two rules of probability

- Two random variables:

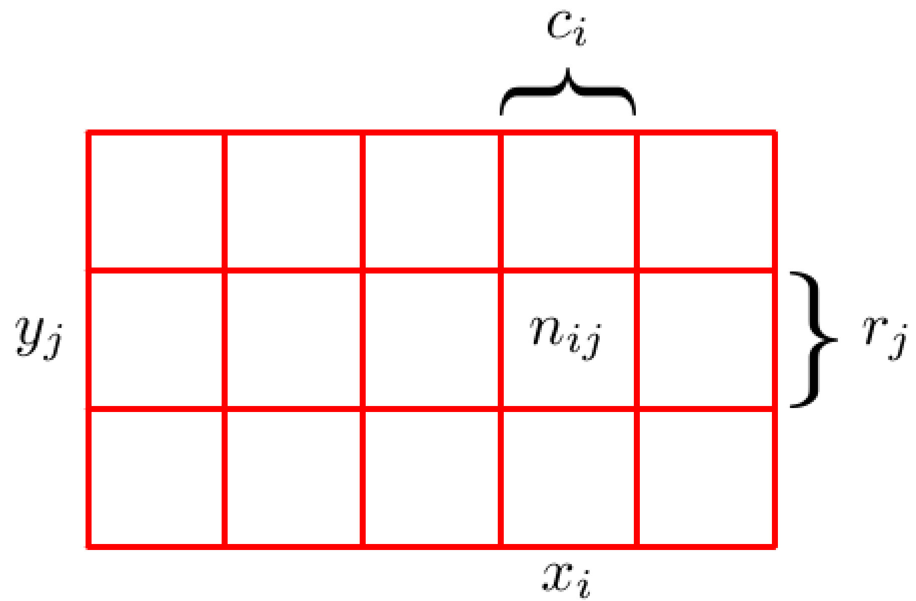
  X takes values $\{x_i\}$ where $i = 1, \ldots, M$

  Y takes values $\{y_j\}$ where $j = 1, \ldots, L$

- Observe outcomes (X, Y) of N samples, with $N \rightarrow \infty$

“joint probability”

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



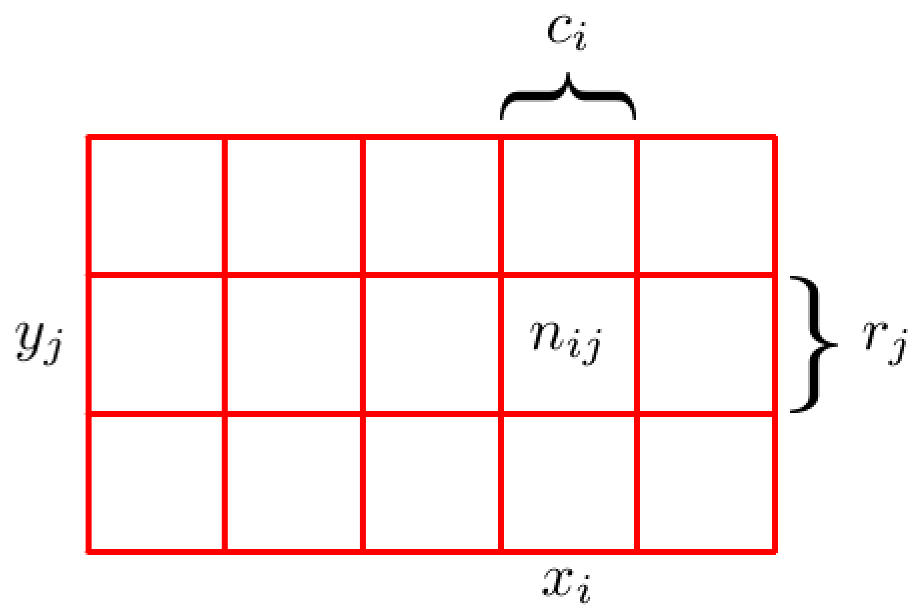
The two rules of probability

- Two random variables:

  X takes values $\{x_i\}$ where $i = 1, \ldots, M$

  Y takes values $\{y_j\}$ where $j = 1, \ldots, L$

- Observe outcomes (X, Y) of N samples, with $N \rightarrow \infty$



"joint probability"

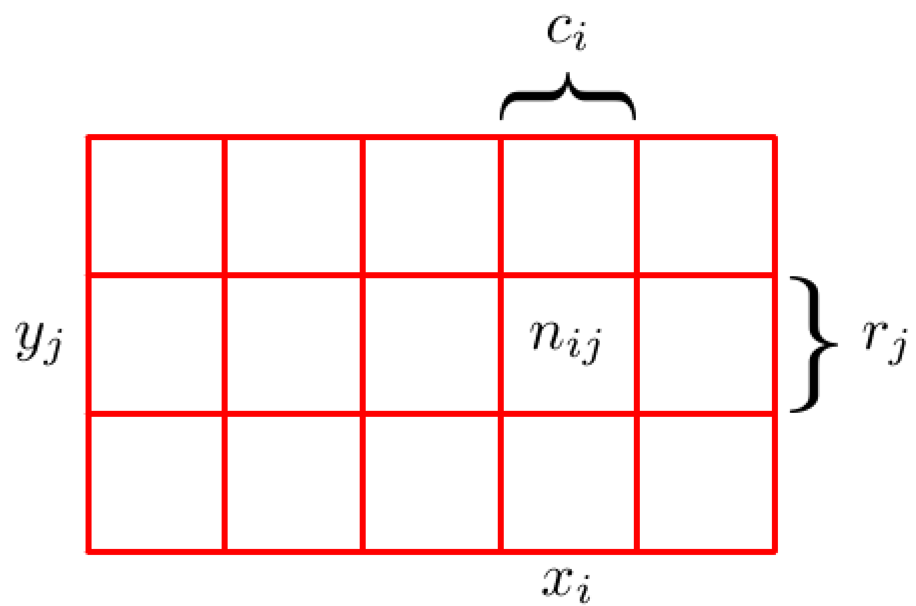
$$
\begin{aligned}
p(X = x_i, Y = y_j) &= \frac{n_{ij}}{N} \\
&= \frac{n_{ij}}{c_i} \cdot \frac{c_i}{N} \\
&= p(Y = y_j | X = x_i) p(X = x_i)
\end{aligned}
$$

The two rules of probability

- Two random variables:

  X takes values $\{x_i\}$ where $i = 1, \ldots, M$

  Y takes values $\{y_j\}$ where $j = 1, \ldots, L$
- Observe outcomes (X, Y) of N samples, with $N \to \infty$



"joint probability"

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

"product rule"

$$= \frac{n_{ij}}{c_i} \cdot \frac{c_i}{N}$$

$$= \quad p(Y = y_j | X = x_i) p(X = x_i)$$

The two rules of probability

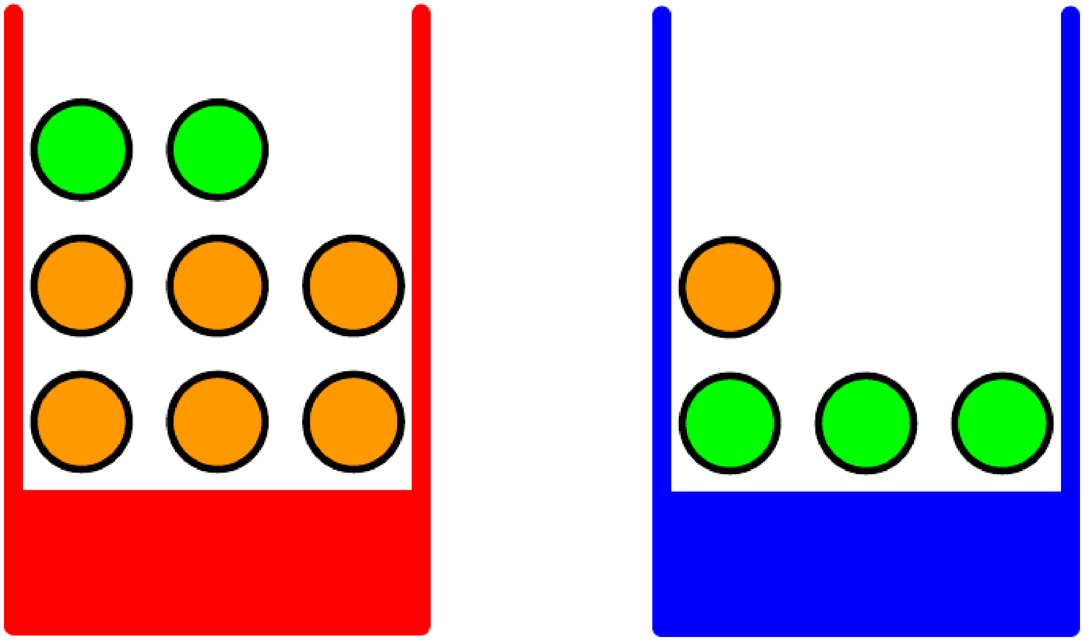
sum rule	$p(X) = \sum_Y p(X, Y)$
product rule	$p(X, Y) = p(Y\|X)p(X)$

– Direct result of product rule: “Bayes' rule”

$$p(Y|X) = \frac{p(X|Y)p(Y)}{p(X)}$$

Apples and oranges

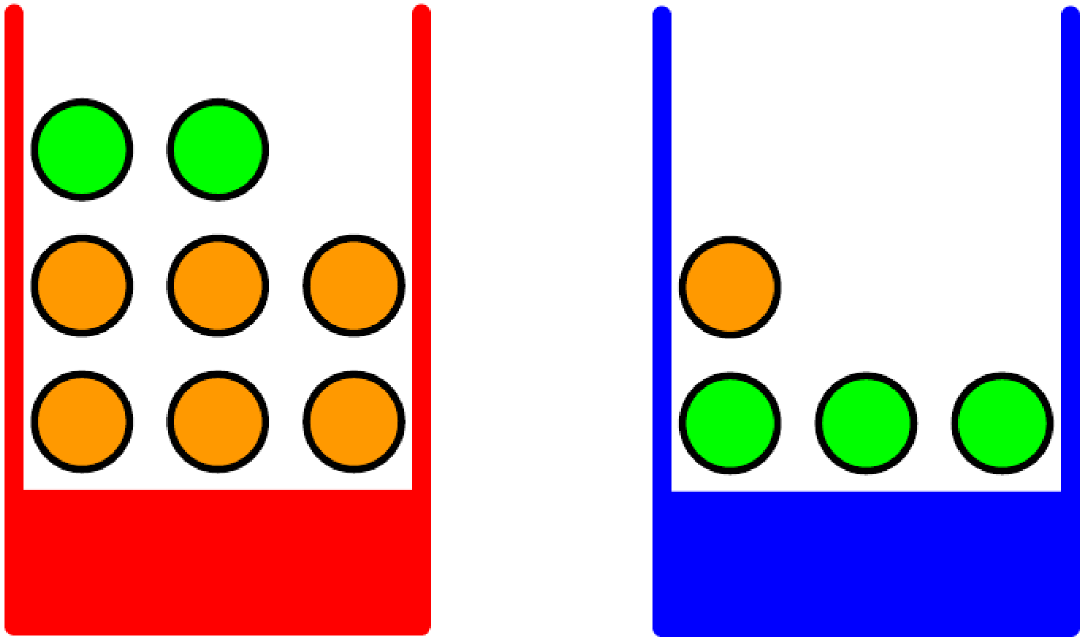
(1) What is the probability we get an apple?



$p(B = r) = 4/10$ $p(B = b) = 6/10$

Apples and oranges

(1) What is the probability we get an apple?

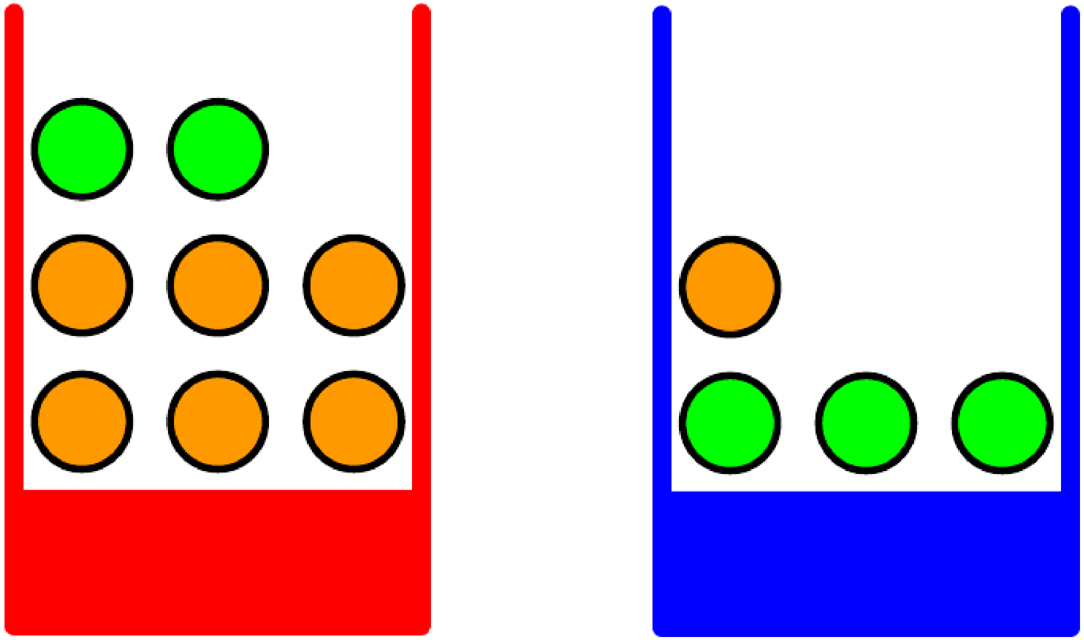


$p(B = r) = 4/10$ $p(B = b) = 6/10$

$$\begin{aligned} p(F = a) &= p(F = a|B = r)p(B = r) + p(F = a|B = b)p(B = b) \\ &= \frac{1}{4} \times \frac{4}{10} + \frac{3}{4} \times \frac{6}{10} = \frac{11}{20} \end{aligned}$$

Apples and oranges

(2) If the fruit we end up with is an orange, what is the probability we had chosen the red box?

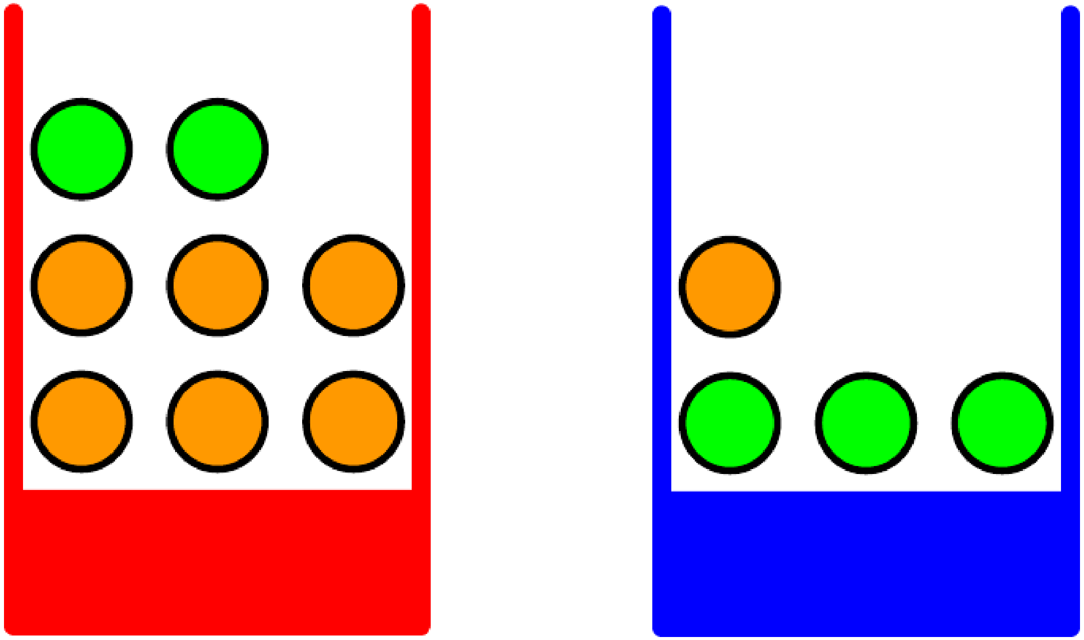


$p(B = r) = 4/10$

$p(B = b) = 6/10$

Apples and oranges

(2) If the fruit we end up with is an orange, what is the probability we had chosen the red box?



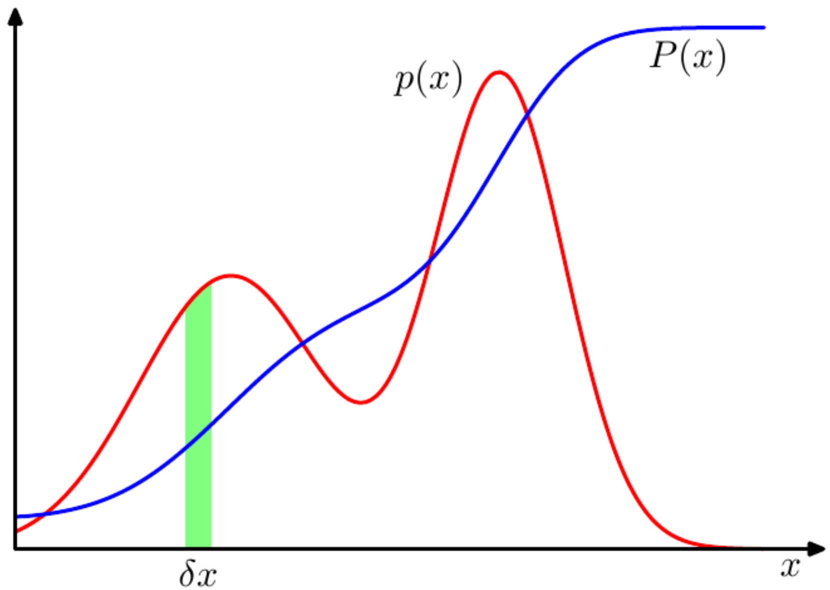
$p(B = r) = 4/10$ $\quad$ $p(B = b) = 6/10$

$$p(B = r|F = o) = \frac{p(F = o|B = r)p(B = r)}{p(F = o)}$$

$$= \frac{3}{4} \times \frac{4}{10} \times \frac{20}{9} = \frac{2}{3}$$

Continuous variables

– $p(x)$ is called the *probability density* over x
 if the probability that x falls in the interval $(x, x+\delta x)$
 is given by $p(x)\delta x$ for $\delta x \to 0$



$$p(x) \geqslant 0$$

$$\int_{-\infty}^{\infty} p(x)\,\mathrm{d}x = 1$$

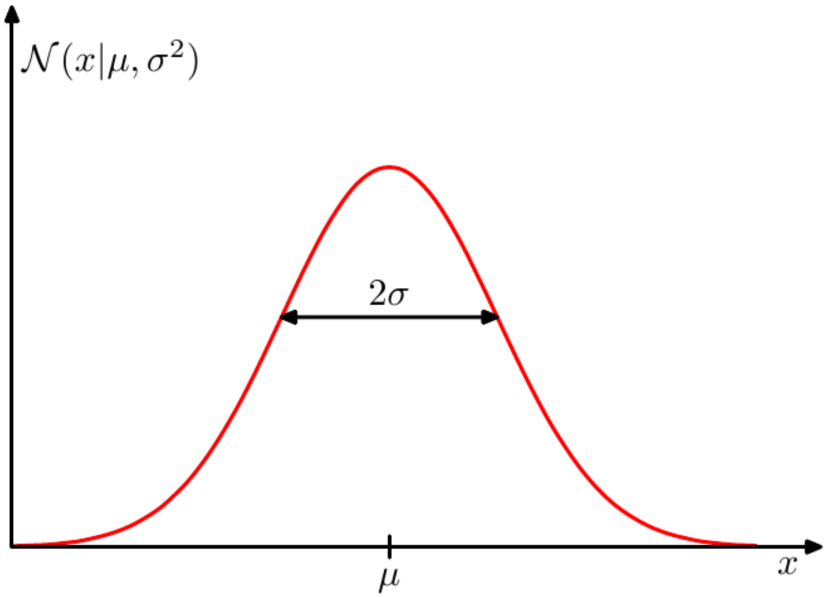
sum rule: $$p(x) = \int p(x,y)\,\mathrm{d}y$$

product rule: $$p(x,y) = p(y|x)p(x).$$

Continuous variables

- Example: Gaussian distribution

$$\mathcal{N}\left(x|\mu,\sigma^2\right) = \frac{1}{(2\pi\sigma^2)^{1/2}} \exp\left\{-\frac{1}{2\sigma^2}(x-\mu)^2\right\}$$



- Multivariate probability density: $p(\mathbf{x}) = p(x_1, \ldots, x_D)$

$$p(\mathbf{x}) \geqslant 0$$

$$\int p(\mathbf{x})\,\mathrm{d}\mathbf{x} = 1$$